

```

1  CCGCAACCCC GACGGCGCCC CAAACGCTGT TGCGCCGCGC GCCCGCCCCA
51  GCCCGGCCTC GCGCTGGTCC CGGTCTCGCC CCGCAGCCCT CGATCTCCCG
101 TGACTTCCTC GGCCAGGCCG CCTGCGCCTC TGGGACCATG TTGCGCTGGC
151 TGGGGGACTT CGCGCTGCCC ACCGCGGCCT GCCAGGACGC GGAGCAGCCG
201 ACCGCTACG AGACCCTCTT CCAGGCACTG GACCGCAATG GGGACGGAGT
251 GGTGGACATC GCGGAGCTGC AGGAGGGGCT CAGGAACCTG GGCATCCCTC
301 TGGGCCAGGA CGCCGAGGAG AAAATTTTTA CTACTGGAGA TGTCAACAAA
351 GATGGGAAGC TGGATTTTGA AGAATTTATG AAGTACCTTA AAGACCATGA
401 GAAGAAAATG AAATTGGCAT TTAAGAGTTT AGACAAAAAT AATGATGGAA
451 AAATTGAGGC TTCAGAAATT GTCCAGTCTC TCCAGACACT GGGTCTGACT
501 ATTTCTGAAC AACAAGCAGA GTTGATTCTT CAAAGCATTG ATGTTGATGG
551 GACAATGACA GTGGACTGGA ATGAATGGAG AGACTACTTC TTATTTAATC
601 CTGTTACAGA CATTGAGGAA ATTATCCGTT TCTGGAAACA TTCTACAGGA
651 ATTGACATAG GGGATAGCTT AACTATTCCA GATGAATTCA CGGAAGACGA
701 AAAAAAATCC GGACAATGGT GGAGGCAGCT TTTGGCAGGA GGCATTGCTG
751 GTGCTGTCTC TCGAACAAGC ACTGCCCTT TGGACCGTCT GAAAATCATG
801 ATGCAGGTTT ACGGTTCAAA ATCAGACAAA ATGAACATAT TTGGTGGCTT
851 TCGACAGATG GTAAAAGAAG GAGGTATCCG CTCGCTTTGG AGGGGAAATG
901 GTACAAACGT CATCAAAATT GCTCCTGAGA CAGCTGTAA ATTCTGGGCA
951 TATGAACAGT ACAAGAAGTT ACTTACTGAA GAAGGACAAA AAATAGGAAC
1001 ATTTGAGAGA TTTATTTCTG GTTCCATGGC TGGAGCAACT GCACAGACTT
1051 TTATATATCC AATGGAGGTT ATGAAAACCA GGCTGGCTGT AGGCAAAACT
1101 GGGCAGTACT CTGGAATATA TGATTGTGCC AAGAAGATTT TGAAACATGA
1151 AGGCTTGGGA GCTTTTACA AAGGCTATGT TCCAATTTA TTAGGTATCA
1201 TACCTTATGC AGGCATAGAT CTTGCTGTGT ATGAGCTCTT GAAGTCCTAT
1251 TGGCTGGATA ATTTTGCAA AGATTCTGTA AACCTGGAG TCATGGTGTT
1301 GCTGGGATGC GGTGCCTTAT CCAGCACCTG TGGTCAGCTG GCCAGCTACC
1351 CATTGGCTTT GGTGAGAACT CGCATGCAGG CTCAAGCCAT GTTAGAAGGT
1401 TCCCCACAGC TGAATATGGT TGGCCTCTTT CGACGAATTA TTCCAAGA
1451 AGGAATACCA GGACTTTACA GAGGCATCAC CCCAACTTC ATGAAGGTGC
1501 TCCCTGCTGT AGGCATCAGT TATGTGGTTT ATGAAAATAT GAAGCAAAT
1551 TTAGGAGTAA CCCAGAAATG ATGTTGCATT TTTTGCTTTA GCCTGATAAT
1601 TGAACTTTC AACAACTCTT GGAGTGACTT TTTCTCCTCG AATTGAAACA
1651 AGTCTATGGC AAAAGAAGCT GCATTTTTTT CACAAAAGGG AAGACGTAA
1701 CAATGGTCAC TTCAAACCTT TGGGCTAAAT TATATGTACA CAGAAATGTT
1751 CAAAATCATA GTTTTAATGT GTTTTGAAAA GGCCACACAA TTATACTTTA
1801 TCTTTTCTTA ATAATCCTGC AAATCTCTGC CCTGAATCCG AAATCTGAAA
1851 ATGTACTGGT TTGAACAAAA TTTGTTTTGT GTGTTAGAGT TATAAATCAT
1901 TAATCTTTAT TTCGGGTGGT TTACGTTTAT GCCAGTTCCT TTATATTTAA
1951 ATTTCTTGTT TTATATATTT TGAATGTCTT TATAGATTTC TTTAAATTTC
2001 CTTATAGAAC CATTAAATAGA AAATCATTAC ATTTAAAATA TACCTTACAG
2051 CAAAAGCATC CAAATAAGTA TAGGGTTTAT GTCCTTATTT TTCTTTCAGC
2101 TGAATACGAA TGAACACAGT GGTGGAATTT CTGAAGGGAA GTGATGAAAT
2151 TATATTTATT TCAGTGGGCA CTTTTCATT TTACCCTGT ACCATTATTT
2201 GGTTCCTGGA GTTATACACT AATTTTCAGT ATATTACTGT TAAATTACCA
2251 ACACAAGGCA ATTTATTTGA AAGATTCCGT TTATCCTGCC ATTGCTTTGA
2301 AAAGCAGCAG GAAACGAAAT TTTTGTACTT GTATCAGCTT CTGCAGAGCA
2351 TCTTTGTTTT CTTTGTCCCT TTGTTTCCTA CTTTGTGAAT CAGATTCCGT
2401 TTTAGTCAGG AAGACTTCTT GGGACCATTC TTAGTAACCT GAAATTTCTT
2451 TTTTAATTGC ATGAAGTGA TTGATCATGA GCAAGTGATG GGCTTTATTT
2501 CTCCTTCACT GGTGAATATC CTTTGAACCT GCTGTTTGCA ATATGGGCAG
2551 CCACAAAGGG GGAGAGATGC CTATTAATC GCGGGGTGT ATGACTTCTG
2601 AAAACATTGG ATACCCTATT TTGAAAAGGG AAAGGCCCAA TTTGGGGAAA
2651 CATATACCAA TGCATGATTT CTG

```

# FEATURES:

```

5'UTR:      1-137
Start Codon: 138
Stop Codon:  1569
3'UTR:      1572

```

**HOMOLOGOUS PROTEINS:**Top BLAST Hits:

	Score	E
CRA 335001098641184 /altid=gi 11360341 /def=pir  T50686 peroxis...	927	0.0
CRA 11000479457833 /altid=gi 6841066 /def=gb AAF28888.1 AF12330...	834	0.0
CRA 18000005183605 /altid=gi 7504235 /def=pir  T22688 hypotheti...	432	e-120
CRA 1000682325160 /altid=gi 7499323 /def=pir  T21074 hypothetic...	377	e-103
CRA 89000000196990 /altid=gi 7294582 /def=gb AAF49922.1  (AE003...	348	9e-95
CRA 150000075553401 /altid=gi 9758252 /def=dbj BAB08751.1  (AB0...	339	5e-92
CRA 335001098657884 /altid=gi 11358611 /def=pir  T49871 peroxis...	330	2e-89
CRA 163000046661776 /altid=gi 10176874 /def=dbj BAB10081.1  (AB...	326	4e-88
CRA 105000014652720 /altid=gi 10798831 /def=dbj BAB16462.1  (AP...	200	3e-50
CRA 335001098655048 /altid=gi 11277065 /def=pir  T47703 Ca-depe...	199	6e-50

BLAST dbEST hits:

gi 10145202 /dataset=dbest /taxon=96...	1108	0.0
gi 1437155 /dataset=dbest /taxon=9606 ...	801	0.0
gi 10333851 /dataset=dbest /taxon=96...	745	0.0
gi 8469752 /dataset=dbest /taxon=960...	363	8e-98
gi 11684041 /dataset=dbest /taxon=96...	307	4e-81

**EXPRESSION INFORMATION FOR MODULATORY USE:**

## library source:

Expression information from BLAST dbEST hits:

gi|10145202 Placenta Choriocarcinoma  
gi|1437155 Retina  
gi|10333851 Uterus leiomyosarcoma  
gi|8469752 Breast  
gi|11684041 Ovary fibrotheoma

Expression information from PCR-based tissue screening panels:

Leukocyte

```

1 MLRWLRDFAL PTAACQDAEQ PTRYETLFQA LDRNGDGVVD IGELQEGLRN
51 LGIPLGQDAE EKIFTTGDVN KDGKLDFFEF MKYLDHEKK MKLAFKSLDK
101 NNDGKIEASE IVQSLQTLGL TISEQQAELI LQSIDVDGTM TVDWNWRDY
151 FLFNPVTDIE EIIRFWKHST GIDIGDSLTI PDEFTEDKK SGQWWRQLLA
201 GGIAGAVSRT STAPLDRDKI MMQVHGSKSD KMNIFGGFRQ MVKEGGIRSL
251 WRNGNTNVIK IAPETAVKFW AYEQYKKLLT EEGQKIGTFE RFISGSMAGA
301 TAQTFIYPME VMKTRLAVGK TGQYSGIYDC AKKILKHEGL GAFYKGYVPN
351 LLGIIPYAGI DLAVYELLKS YWLDNFAKDS VNPGVMVLLG CGALSSTCGQ
401 LASYPLALVR TRMQAQAMLE GSPQLNMVGL FRRIISKEGI PGLYRGITPN
451 FMKVLPAVGI SYVVYENMKQ TLGVTQK

```

# **FEATURES:**

## **Functional domains and key regions:**

[1] PDOC00001 PS00001 ASN\_GLYCOSYLATION  
N-glycosylation site

254-257 NGTN

[2] PDOC00005 PS00005 PKC\_PHOSPHO\_SITE  
Protein kinase C phosphorylation site

Number of matches: 2

1	229-231	SDK
2	475-477	TQK

[3] PDOC00006 PS00006 CK2\_PHOSPHO\_SITE  
Casein kinase II phosphorylation site

Number of matches: 8

1	22-25	TRYE
2	65-68	TTGD
3	121-124	TISE
4	157-160	TDIE
5	170-173	TGID
6	179-182	TIPD
7	185-188	TEDE
8	227-230	SKSD

[4] PDOC00008 PS00008 MYRISTYL  
N-myristoylation site

Number of matches: 16

1	52-57	GIPLGQ
2	119-124	GLTISE
3	171-176	GIDIGD
4	201-206	GGIAGA
5	202-207	GIAGAV
6	245-250	GGIRSL
7	253-258	GNGTNV
8	283-288	GQKIGT
9	295-300	GSMAGA
10	322-327	GQYSGI
11	326-331	GIYDCA
12	359-364	GIDLAV
13	392-397	GALSST
14	399-404	GQLASY
15	442-447	GLYRGI
16	446-451	GITPNF

[5] PDOC00018 PS00018 EF\_HAND  
EF-hand calcium-binding domain

Number of matches: 3

1	32-44	DRNGDGVVDIGEL
2	68-80	DVNKDGKLDFFEEF
3	99-111	DKNNDGKIEASEI

**Membrane spanning structure and domains:**

Helix	Begin	End	Score	Certainty
1	292	312	1.053	Certain
2	345	365	0.613	Putative
3	381	401	1.544	Certain
4	446	466	0.733	Putative

**BLAST Alignment to Top Hit:**

>CRA|335001098641184 /altid=gi|11360341 /def=pir||T50686 peroxisomal  
Ca-dependent solute carrier [imported] - rabbit  
/org=rabbit /taxon=9986 /dataset=nraa /length=475  
Length = 475

Score = 927 bits (2371), Expect = 0.0  
Identities = 454/477 (95%), Positives = 466/477 (97%), Gaps = 2/477 (0%)

Query: 1 MLRWLRDFALPTAACQDAEQPTRYETLFQALDRNGDGVVDIGELQEGLRNLGIPLGQDAE 60  
MLRWLR F LPTAACQ AE PTRYETLFQALDRNGDGVVDI ELQEGL++LGIPLGQDAE  
Sbjct: 1 MLRWLRGFLVPTAACQGAEPTRYETLFQALDRNGDGVVDIRELQEGLKSLGIPLGQDAE 60

Query: 61 EKIFTTGDVNKDGKLDFFEEFMKYLDHEKKMKLAFKSLDKNNDGKIEASEIVQSLQTLGL 120  
EKIFTTGDVNKDGKLDFFEEFMKYLDHEKKMKLAFKSLDKNNDGKIEASEIVQSLQTLGL  
Sbjct: 61 EKIFTTGDVNKDGKLDFFEEFMKYLDHEKKMKLAFKSLDKNNDGKIEASEIVQSLQTLGL 120

Query: 121 TISEQQAELILQSIDVDGTMVTDWNEWDRDYFLFNPVTDIEEIIIRFWKHSTGIDIGDSLTI 180  
TISEQQAELILQSID DGTMTVDWNEWDRDYFLFNPV DIEEIIIRFWKHSTGIDIGDSLTI  
Sbjct: 121 TISEQQAELILQSIDADGTMVTDWNEWDRDYFLFNPVADIEEIIIRFWKHSTGIDIGDSLTI 180

Query: 181 PDEFTTEDEKKSGQWWRQLLAGGIAGAVSRTSTAPLDRKIMMQVHGSKSDKMNIFFGGFRQ 240  
PDEFTTE+E+KSGQWWRQLLAGGIAGAVSRTSTAPLDRK+MMQVHGSKS MNIFGGFRQ  
Sbjct: 181 PDEFTTEERKSGQWWRQLLAGGIAGAVSRTSTAPLDRKVMQVHGSKS--MNIFGGFRQ 238

Query: 241 MVKEGGIRSLWRGNGTNVIKIAPETAVKFWAYEQYKKLLTEEGQKIGTFERFISGSMAGA 300  
M+KEGG+RSLWRGNGTNVIKIAPETAVKFW YEYKKLLTEEGQKIGTFERFISGSMAGA  
Sbjct: 239 MIKEGGVRSWRGNGTNVIKIAPETAVKFWVEYQYKKLLTEEGQKIGTFERFISGSMAGA 298

Query: 301 TAQTFIYPMEVMKTRLAVGKTGQYSGIYDCAKKILKHEGLGAFYKGYVPNLLGIIPYAGI 360  
TAQTFIYPMEVMKTRLAVGKTGQYSGIYDCAKKILK+EG GAFYKGYVPNLLGIIPYAGI  
Sbjct: 299 TAQTFIYPMEVMKTRLAVGKTGQYSGIYDCAKKILKYEGFGAFYKGYVPNLLGIIPYAGI 358

Query: 361 DLAVYELLKSYWLDNFAKDSVNPVGMVLLGCGALSSTCGQLASYPLALVRTRMQAQAMLE 420  
DLAVYELLKS+WLDNFAKDSVNPV+VLLGCGALSSTCGQLASYPLALVRTRMQAQAMLE  
Sbjct: 359 DLAVYELLKSHWLDNFAKDSVNPVGLVLLGCGALSSTCGQLASYPLALVRTRMQAQAMLE 418

Query: 421 GSPQLNMVGLFRRIISKEGIPGLYRGITPNFMKVLPVAVGISYVVYENMKQTLGVTQK 477  
G+PQLNMVGLFRRIISKEG+PGLYRGITPNFMKVLPVAVGISYVVYENMKQTLGVTQK  
Sbjct: 419 GAPQLNMVGLFRRIISKEGLPGLYRGITPNFMKVLPVAVGISYVVYENMKQTLGVTQK 475

>CRA|11000479457833 /altid=gi|6841066 /def=gb|AAF28888.1|AF123303\_1  
(AF123303) calcium-binding transporter [Homo sapiens]  
/org=Homo sapiens /taxon=9606 /dataset=nraa /length=411  
Length = 411

Score = 834 bits (2132), Expect = 0.0  
Identities = 409/410 (99%), Positives = 409/410 (99%)

Query: 8 FALPTAACQDAEQPTRYETLFQALDRNGDGVVDIGELQEGLRNLGIPLGQDAEEKIFTTG 67  
F LPTAACQDAEQPTRYETLFQALDRNGDGVVDIGELQEGLRNLGIPLGQDAEEKIFTTG  
Sbjct: 1 FVLPTAACQDAEQPTRYETLFQALDRNGDGVVDIGELQEGLRNLGIPLGQDAEEKIFTTG 60

Query: 68 DVNKGDKLDFFEEFMKYLDHEKKMKLAFKSLDKNNDGKIEASEIVQSLQTLGLTISEQQA 127  
DVNKGDKLDFFEEFMKYLDHEKKMKLAFKSLDKNNDGKIEASEIVQSLQTLGLTISEQQA  
Sbjct: 61 DVNKGDKLDFFEEFMKYLDHEKKMKLAFKSLDKNNDGKIEASEIVQSLQTLGLTISEQQA 120

Query: 128 ELILQSIDVDGTMVTDWNEWDRDYFLFNPVTDIEEIIIRFWKHSTGIDIGDSLTIIPDEFTED 187  
ELILQSIDVDGTMVTDWNEWDRDYFLFNPVTDIEEIIIRFWKHSTGIDIGDSLTIIPDEFTED  
Sbjct: 121 ELILQSIDVDGTMVTDWNEWDRDYFLFNPVTDIEEIIIRFWKHSTGIDIGDSLTIIPDEFTED 180

Query: 188 EKKSGQWWRQLLAGGIAGAVSRTSTAPLDRKIMMQVHGSKSDKMNIFFGGFRQMVKEGGI 247

EKKSGQWWRQLLAGGIAGAVSRTSTAPLDRLKIMMQVHGSKSDKMNIFFGGFRQMVKEGGI  
 Sbjct: 181 EKKSGQWWRQLLAGGIAGAVSRTSTAPLDRLKIMMQVHGSKSDKMNIFFGGFRQMVKEGGI 240  
 Query: 248 RSLWRGNGTNVIKIAPETAVKFWAYEQYKKLLTEEGQKIGTFERFISGSMAGATAQTFIY 307  
 RSLWRGNGTNVIKIAPETAVKFWAYEQYKKLLTEEGQKIGTFERFISGSMAGATAQTFIY  
 Sbjct: 241 RSLWRGNGTNVIKIAPETAVKFWAYEQYKKLLTEEGQKIGTFERFISGSMAGATAQTFIY 300  
 Query: 308 PMEVMKTRLAVGKTGQYSGIYDCAKKILKHEGLGAFYKGYVPNLLGIIPYAGIDLAVYEL 367  
 PMEVMKTRLAVGKTGQYSGIYDCAKKILKHEGLGAFYKGYVPNLLGIIPYAGIDLAVYEL  
 Sbjct: 301 PMEVMKTRLAVGKTGQYSGIYDCAKKILKHEGLGAFYKGYVPNLLGIIPYAGIDLAVYEL 360  
 Query: 368 LKSYWLDNFAKDSVNPVGVMVLLGCGALSSTCGQLASYPLALVRTRMQAQA 417  
 LKSYWLDNFAKDSVNPVGVMVLLGCGALSSTCGQLASYPLALVRTRMQAQA  
 Sbjct: 361 LKSYWLDNFAKDSVNPVGVMVLLGCGALSSTCGQLASYPLALVRTRMQAQA 410

Score = 80.0 bits (194), Expect = 6e-14  
 Identities = 80/388 (20%), Positives = 156/388 (39%), Gaps = 59/388 (15%)

Query: 95 FKSLDKNNDGKIEASEIVQSLQTLGLTISEQQAEILILQSIDV--DGTMTVDWNEWDRDYFL 152  
 F++LD+N DG ++ E+ + L+ LG+ + + E I + DV DG +  
 Sbjct: 21 FQALDRNGDGVVDIGELQEGLRNLGIPLGQDAEEKIFTTGDVNKDGLK----- 68  
 Query: 153 FNPVTDIEEIIIRFWKHSTGIDIGDSLTPDEFTDEKKSGQWWRQLLAGGIAGAVSRTST 212  
 D EE +++ K + EKK ++ L +  
 Sbjct: 69 -----DFEEFMKYLK-----DHEKKMKLAFKSLDKNNDGKIEASEIV 105  
 Query: 213 APLDRLKIMMQVHGSKSDKMNIFFGGFRQMVKEGGIRSLWRGNGTNVIKIAPETAVKFWAY 272  
 L L + + ++ +I V R + N I E ++FW +  
 Sbjct: 106 QSLQTLGLTISEQQAEILILQSIDVDGTMTVDWNEWDRDYFLFNPVTDI----EEIIRFWKH 161  
 Query: 273 EQYKKL-----LTEEGQKIGTFER-FISGSMAGATAQTFIYPMEVMKTRLAV-GKT 321  
 + TE+ +K G + R ++G +AGA ++T P++ +K + V G  
 Sbjct: 162 STGIDIGDSLTPDEFTDEKKSGQWWRQLLAGGIAGAVSRTSTAPLDRLKIMMQVHGSK 221  
 Query: 322 GQYSGIYDCAKKILKHEGLGAFYKGYVPNLLGIIPYAGIDLAVYELLKSYWLDNFAKDSV 381  
 I+ ++++K G+ + ++G N++ I P + YE K ++  
 Sbjct: 222 SDKMNIFFGGFRQMVKEGGIRSLWRGNGTNVIKIAPETAVKFWAYEQYKKL----LTEEGQ 277  
 Query: 382 NPGVMVLLGCGALSSTCGQLASYPLALVRTRMQAQAAMLEGSPQLNMVGLFRRIISKEGIP 441  
 G G+++ Q YP+ +++TR+ A+ + + ++I+ EG+  
 Sbjct: 278 KIGTFERFISGSMAGATAQTFIYPMEVMKTRL---AVGKTGQYSGIYDCAKKILKHEGLG 334  
 Query: 442 GLYRGITPNFMKVLPAVGISYVYENMK 469  
 Y+G PN + ++P GI VYE +K  
 Sbjct: 335 AFYKGYVPNLLGIIPYAGIDLAVYELLK 362

#### Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00153	Mitochondrial carrier proteins	305.4	3e-88	1
PF00036	EF hand	50.7	1.7e-12	3
PF00404	Dockerin domain type I	9.7	0.26	1
PF01978	Protein of unknown function	2.7	9.5	1

#### Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00036	1/3	27	51 ..	5	29 .]	18.7	0.002
PF00404	1/1	67	85 ..	1	22 []	9.7	0.26
PF00036	2/3	61	87 ..	3	29 .]	19.7	0.001
PF00036	3/3	90	118 ..	1	29 []	17.2	0.0051
PF01978	1/1	110	121 ..	1	13 [.	2.7	9.5
PF00153	1/1	193	472 ..	1	313 []	305.4	3e-88

```

1  AACCCATGTT AGTGTGCAGT TCTGCTGGCA CACACATGCA GTTGTGTAAC
51  CACTACCACC AAAAGCAAGA TGTAATAATAG CTCCATCACC CCCACAAGCC
101 TTCTGATGCT CTTTTGTCAT CAATTCCTT CCCGCTAGTC ACAACTGGTA
151 ACTACTGATT TGTTTTCTGT CCCTATAGTT TTGCCTTTTC CAGAATGTCA
201 TTGTTGACAG GTATCAGTAA TTCATTCCCT TTTATTGCTA ATTACTATCT
251 CACTGTATGA ATGCAACACA GGTTGTTTAC CAGTTCACCC GTTAAAGAAC
301 ATTTTGTTC TGCGCTTGAC AGTTATGAAT AGAACTGCTA TAAACCCTCA
351 AGTAAAAGTT TTGGTGTGAA GATAATTTTC TCAGCAAAAA CGCTGACAGG
401 TAATTTTCT AAGTATTACT TTTTAAAAA AGTAAAATAG CCTGTAGCCC
451 CAGCTACTCA GGAGGCTGAG GCAGGAGAAT AGCTTGAACC CAGGAGGCGG
501 AGGTTGCAGT GAGTTGAGAT TGTGCCACTG CATTCCAGCC TGGGCGACAG
551 AGCTAGACTG TCTCAAAGAA AAAAAAATAA AATAACAAAT AAATAAAAG
601 TAAAAAGAAA GCATGTAAGT GTAAGATGAC TAGTTCAAGC AACCTCTCTT
651 CAAGTACAGA GTATTAGAG TAGAGATTAA AAGAGGTTT CAAGGACAGA
701 GAAAATTGTA AGTTTGAAGG CAGTTCCAAA GGAAGGCAAT GATTCTTAAT
751 AAGACTGGAA GTTGAAGTA ATATAAAAG ATAAATCAGT TTCAAGATGA
801 TTTTACTAAG CAGGCAGCCC TTAATTTACA AATTCTAGAT TCATACATAT
851 CTTAAACATA CAAAATGATA TGAGGAGAGG TAAGTTCAGG GTCTGAGTTC
901 CTGGCTGTTG TTGGAAGTGA TTTCTGTGTA GTGATTGAGA AGATGTGAGA
951 CACCCTAATT TACAAGTACA GAGGTATCTT CTTTTCTGCA AACAGCAGTA
1001 CAACAATAGT TCCTCTTACG CAGCTGTGAA TGAACAGGAT TATTACAATT
1051 AATGATATCT CATTTGATTG GCGCCTTAGA GAATTAAGAC CTTTCACACC
1101 TAATATACAA CTTTGTGTG AAGGCAGATA TTTATATTCT CATTTTACTG
1151 ATGAGAGACT ACCCGGAGAC GCTATGTCAC ACCTGAAGGA TTAGGTACTT
1201 TCTCTGTAA GTCCAATGTT CCTTCCGTTA TTCCATGCTA GGCAGTAATA
1251 AGTTCTGTCT TGCCTGAGTA ATAAGCTCCA AACCTCGGAA CTGCACCCAT
1301 CTTGAGAAGG AGGAGGGCGC TGTGGTTTTT TCTGATAAGT GCAGCTGGCA
1351 GACACTCTAT ACGCTTAATC ACGGGCAAAT CCTACCTAAG CTGCCTACCA
1401 AACTAGTCCT TCTTTTCCCC GTTGCCACAG CAGATGGCTG TTGATCTTTT
1451 CTGCAACAAA TCCAGGAGTT TCTCCTTTTT GTTTTATAAT TGCTCCAATA
1501 GATGCTTTAG GATTTAACTC TCTGCTTTTT AAAGCAGAAT CGCCATCCCA
1551 GGTGTGCAAC CACGAAAAAA TTAGACATCC GTGAGAGACA ATGCCCTCCA
1601 TGCCCCAGTT TCCAGGCAGA GAGAAGCAGC TCTGGGCTGA CCGCCAAGGC
1651 TCCGGCCCCG GAGGGTCTTT AAGTGGAGTA ACCAGTCTTC AAGACCCCGC
1701 TCCCAAGCCA CCGACGCGCT GACGCTGCAG CCCTGGACCT GCTGGGGGCC
1751 TCTTCCTCGG ACCCGCATGC TGACAGCGGG ACTGGCAACT GGGCAGAGGT
1801 CGACCCCGGG TCCGCACAGC ACCTCCCGAG ACCCAGCTCC CAGCTCCCTC
1851 ACTTCCGGCT CTCTGGAGGC GGGCCCGGCC AGTGCCCGCG AGGCCAGCGC
1901 GCGGAGCTCC TCCCCAGCAG CCGCGGGACG GCCACACCCT GCGCGCCGCG
1951 CGGGCTCGGG TGGGGTCTCC GCTCCTGCGC CCTGCGCGCC GCAGCCGCAC
2001 CCCCACGGC GCCCAAACG CTGTTGCGCC GCGCGCCCCG CCCAGCCCGG
2051 CCTCGCGCTG GTCCCGTCT CGCCCCGAG CCCTCGATCT CCCGTGACTT
2101 CCTCGGCCAG GCCGCTGCG CCTCTGGGAC CATGTTGCGC TGGCTGCGGG
2151 ACTTCGTGCT GCCCACCGC GCCTGCCAGG ACGCGGAGCA GCCGACGCGC
2201 TACGAGACCC TCTTCCAGGC ACTGGACCGC AATGGGGACG GAGTGGTGGA
2251 CATCGCGCAG TCTCAGGAGG GGCTCAGGAA CCTGGGCATC CCTCTGGGCC
2301 AGGACGCCGA GGAGGTGGGT CGCCGCCGGG GCGCCGCTG AGCGTAGGGA
2351 GGGCTGCGGG CGCTGGGGAC ACTGCGAGGA CCGAGGAGGG CGGCGGCTTG
2401 AGGCGTTGCC AGGAGAGGAA GGAGGAACTG TGGCGCCAG CGCTCCGGTG
2451 GCTTCAGAAA CTCGGGCGTG GGGCCGCGAC CGGCGACCCC GGTAAACAGAA
2501 GTGGGTCTAT ATACGAAAGT CTAATGGTAT TTGTCCAGAT AAAATGAGTG
2551 TTGTGGACAC TCTGGCCAC GGGCACTGTT AAATTTTAA GACACTTTTG
2601 TCCTGAATCC ATCCAGGTT CTTTGTTTTC TGTTTTAATA CCTTGCAGAC
2651 ATGTAATCCG TTTTAGCTGT CAGACTTCAG TGGGTCCAA GTTTGTGATA
2701 AAGCGCACCA CATTCGATCT CTTTCGAAGC TGCTTTGTTA CAGCAGCTAT
2751 GTGTATTGTC TACTGTTTGA AAAGTGTTC AAAACCAATC GCGTGTTCCT
2801 CCCACTTCCT GTTGAGAAGG AATGGCGGCA TTCCATTGTT TAAGACATTC
2851 CTAGGTTAAT GCCTAGGTA CATAAATTGA TCTGAAGGTT TGACTTGACC
2901 TCGGACTGAG CAATTTTATT TTCTCTGAGT CATCTTAAGT GTGCCCTGA
2951 ACTTCTGCCC CTTTAGTAGG GTGGAGATAT GTGGAACCTC TCCAACCCTG
3001 TTGAAGCGTT CCCTGACACT GGCATTCTCT TATCCAAAGA GGGAAAGTGA
3051 TTAGGTTACT ATGAGGGCCA ACAACTGTTA TATAGTTATA TTTCACTTCT
3101 CTTTAAATGT CTTTGGTAGT TATAGGCCCTC TTCAGTTTAC TGTTTCTTCT

```

FIGURE 3, page 1 of 42

3151 AGAGTCAGAT TTAGTAAGTT ACAATTTTTT TTGAAACTGC CTGTTCTGTC  
3201 CAAGGTTTCAT AATACTCACC GATGATTTTA TAACACTTCT GACTGAATCT  
3251 GTAGGTAGGT TCTCTATTTC ATTCCTCATA TCTATCCTTT TCTCCCCCTC  
3301 AATCTTGCCA AAGTTTTGTG TATTTTATTC ATACTTTGAA GGAACCAACT  
3351 TTTGGTACTT TGTGCTGATT GTCCAGAAA TGGCCAGTT GGAGTTCCCC  
3401 ACCATGTCCA ATCATTGGCT GGAAGCAGCC CAGGAAAGGG ACGACCTTGC  
3451 TGCAGTGCAT CAGCAGATGC CAGGGTTAGA GGCTAGAGAG TGAAGTCAA  
3501 CTGTGTTCCCT CACAGTAGGT GCCTTTGAAG GGAGATCTCA GTGGTACAAC  
3551 TCCATGGTCC CTACAATATA CAAAAGCTCT TTGGAGTGCT CAATGATTTT  
3601 TAAGATTGTA AAGGGATCCT GAGATCAAAA AGCTTGAGAA TTGCTGCTGT  
3651 ATCACCATTT TTACGTAAC GCATCATATT CTGTTATATG TTGTGTGCT  
3701 AGTATATGTT ACCAATTCTT TTTAAATCAC CTTTACTTTT ATTGATAGTT  
3751 TAAAAACGAT TGTAAGTGAA ATTGCAATGG ATGTCCTTTG TATTCATTTT  
3801 CTCATTCTGG TCCAGTTACT TTCGTAGGAT AAATTTTGAG GAGTGGACAT  
3851 TGCTGAGTCT GAAGGTAACA CACATTTTAA ACTGGGATAC GTATTGCCTT  
3901 TCGGAAACCT TAGACCCATT TTCACTCTTT TGAAGTACAG TGCTTGCTTC  
3951 TCCACATCCT CGCTCATPCA GGGTATCAGT CTTTGTAAG TCTCCTATTC  
4001 TGCAGGTGAA ATTCCTTTTC ATTTCTGTC TTAGTCCATT TAGTGTGCT  
4051 ATAGTGGAAAT ATCTGAGACA GGGTAATTTA TAAAGAAAAG ACATTTATTT  
4101 AGCTCACAGT TCCGAGGCT GGAAGTTA AGAAGCGTGG TGCTGGCATC  
4151 TGCTGGACTC CTGGGGAGGG CTTTCTGCT GTGTCACAAC ATGGTGGAAA  
4201 GTCAAAGTGG AAGTGGACAT GTGTGAAGAA GCAAAATCCG AGGGGTGTCC  
4251 TGGCTTTATA GCAACCCAGC CTCGAGGGAA CTGATCCATT ACTGAGGGAA  
4301 CTAATTCAGT CTCATGAGAG AGAGAACTCA CTCACTACTG CAAGAATGAC  
4351 ACCAAGCCAT TCATGAGGGA TCTGCCTCCG TAACCCTGAC ACCTCCTGCT  
4401 AGGTCCCTCC TCCCAACACG GCCACATCAG GGATCAGACT TCAACATGAG  
4451 TTTTGTGGG GACAAACAAA ACGTAGCACT TGCTTTGCCT TTTGGTTCTA  
4501 TTCACATCCT CCACAGGATT GCATTATGCC TACCCATTG GTGAGGCGAG  
4551 TCTTCTTTAA TTGGTTTACT GATTCAAATG CTACCCCTCT CCAGAGACAT  
4601 CCTCACAGAC ACACCCAGAA ATCATGTTTT ACCAGTTATC TGGGCATCCC  
4651 TTAGTCCAGA CGAGTTGATA CATAAAATTA ACCATCACAC ATGGGATAGA  
4701 ATTAGGATTA CACAGTCAAC CTTTATGGGA GAAAATTTCA GAGGCATGTC  
4751 AGGGGTTTAT GTAAATGTCAA GGAGTGAGGA CATTGGCTAC TTGAGCATAG  
4801 AAATGAGAAC TGTGGGGTGA CTCTTCGGTG GAAAGTTTCA AGGTAGTAGT  
4851 TTGTATCTAA GCCAAATACT CAGCTTGAAG CAAAATCTCT ATAAATTTTC  
4901 ATCTGATTTG ATCTCATCTC CGTGTTTCCA AGCATTGTGA ATGAATTGAG  
4951 CATTTAGAAG AGAACAAATT TCTGTTTAA TTTCTTTAGA TTTTAGATGG  
5001 AAAGAATGTA GAAATAAGAG TAGAATGTAG AAATAGGTAT AAAGAATATA  
5051 ATAGCTAACC ATTACTAAGT GTTCCAGAAT TATCCAGGGA AGAGAAAAGA  
5101 ATTCAGGCA AGTCTGAGA CAAAATTAAG AACCAATTGG AAGTGAAAGC  
5151 GCTACATTTT TTTTCTCTGG TATGACCTTT CTTTCTATA TGTTCCAAAT  
5201 CTCCTCACTA TGAAATTAGT GAAAAATTAA AGTTAAAAAT TAGAGAAAAT  
5251 TCACATTAAG TTCTCCTAGG ACTCAGTAGT ATAAGGGTAT AGACTGAGAG  
5301 TAGAATGTAG TGTGAGAACA AGGAGATACA GTATTTAACC ATTACTAAT  
5351 CTCTTATACT TGTCTAGTAA TCCTATTTCC TTTTAAAAGT CTTCAGTTAT  
5401 TTTCTCTTTA CGCACCTCCT TCTCCCTCTT GTCTTCTCTC TTCTACCCCC  
5451 ATCTTTCTTC CTGTGGAGCC TTCATGAATG GGATTAGTGC TTGTATAAAA  
5501 GTGACCTGGA AGACCTTCTT TGCCCTTCC ACCATGTGAG GACACAGTGA  
5551 GAAAACAGTG GTCCATGGAA CCGGAAAGTG GGTCTCACT AGACAGTAAA  
5601 TCTCCTAGCA CTTGATCTA GGACTTCCAG TGTCTGGAAC TGCAAGAAAT  
5651 CAATGCTTAT TGTTTAAAGT AGCCAGTAGT ATTTTGTGCA TAGCAGCCCA  
5701 GTTGGACTAG GACAATTACC AAGAGCAAGA AGGGAAGCAG CAAGCTACAA  
5751 GAGAGTTCCG TCCTTGGTGT AAATTGACCG TGTAATCCTT GTCAAGTTTG  
5801 AGCCTTACTG GAGCTTTACT TTCTTATTCT TAAAATGCAG ATATCTTGCC  
5851 TGCATCCTGG ACAGAGCTTT TAACAAGGTC ATATGTTGCA GAATATGAAA  
5901 GTTCATGTTA AAAAACCTTT TAAATGTGG TATCCCATTT ACTAGCTGGT  
5951 GAACTTCTTG AGGAACCTCT GTGCCATGG GTATGAAGTG TATGCTGAAT  
6001 GATCACCCAA TGTAGAGGA GTGGGTGGAC TGGTAACCTG ATTTAAGGGC  
6051 CATTCTAACT CTTACATTCT ATGATTTTTT TAATTCTGTC TTTAAGTTTT  
6101 TACATTTACA ATCACAGAAA AAATAGTCAC ATAGAAGAAT AGTAGCTTAG  
6151 CAAATGTTTA TTGCATTGAG TGGAAATCAGG ATTTCACTCC ATTAAGTAAT  
6201 TCCTCTGTTA ACAAAGAGGG TTCATTCAT TTTTATTTCA TTAATATTGC  
6251 TTTTTTTTTT TTTTTTCTGG AGACAGAATC TTGCTCTATC ACCAAGGCTG

FIGURE 3, page 2 of 42



6301 GAGTGCAGTG GTGCGATCTC GGCTCACTGC AGCCTCTGCT TCCTGGATTTC  
6351 AAGCGATTCT TGTGCCTCAG CCTCCCAAGC AGCTGAGATT ACAGGCACAT  
6401 GCCACCACAC CTGGTTAACT TTTGTATTTT CTAGTAGAGA TGGGATTTTG  
6451 CCATGTTGGT CAGGCTGGTC TTGAATTCCCT GGCTCTAGT GATCTGCCTG  
6501 CCTCTGCCCTC TGAAGTGCT AAGATTACAG GCATGAGCTA CCATGGCCAG  
6551 CCCATTTCCCT TAATATTTTA ATTGTCAGAC ATGTTATGGT TTCTGGCACA  
6601 ATATTAAGAA GACATGATAT GAAATCACAG GGTGAATTTT AGGGCATCAC  
6651 AACAGAAAGA TTATGGTATA AGAAAAACAA TGAATTTCCA ACTACATTTT  
6701 TGTCAAATGT TCTAAAATAT ATAAAATCTG TATCTTTTGT GTTCTCTCCT  
6751 GATTTATATT CTAAATTTGA TGTTATCCTT CTCTGCAGAA ATAAAGTGTC  
6801 TGAAAGAATG AAAAAATGG AAGAATCTT TAGTAAGGTA TAAAATACCC  
6851 TTTCTATCTT TGTAGCATTC TAAGCCTTTT GTCACTTTT CAAACTCCCA  
6901 ACATGCCATA TTCCCTGACT AGGCCACAGC CATGTACATT GATCCCTTTA  
6951 TTTTCTTCTC TCTGCCTGAG ATTTCTCTCA TTCCCCCTTC TCTGCCTGGT  
7001 ATATGATTGC CCATTGTTTA AGGCCCAAC TCACCTTTAT AATCTTCCTA  
7051 GCCCACTTTC TTTATCGGTA TTCCAGAAAA AACAAAAGAA GCTTCCACAA  
7101 GACAACATTC TGTAATACAC TGCTTAACTT CTTTGTACCC TGCTGAGTTC  
7151 AAAAATCTTA TCTTTTAAAG GATTGAATGG AGTCCACCAA GGTATCTATA  
7201 TTTGACAGGA TTTATGAAAA CAAAAGGATT TGGTGAGAAA GTTTGAAGCC  
7251 TAACTCTGAA ACGTGGATCA TAGTGTTTAC TACACATTAA CTGTTTTAGT  
7301 GGATGTAATA GTTATTATTA TAGGCTGTGG AATCAGAAAC GGGTTCAAAT  
7351 GTTTTCACCG CTGCTAGAC TGTGGCCTTG GGCATGTTAT TTAATGCCTG  
7401 GAGGCCTCAA ATGTTAACTA GGAATGGTAA GACCTACCCA GTAACCTAGC  
7451 ATAAATAGTA AATTCATTCA TTTAATGTTT TCAAACAGTG CCAGACATTG  
7501 TTTAATGAAC TGGGGATATA GTGGTGAACA AACTGACAG CGTCTTCAT  
7551 TGTATTCTCA AAACCTCCC TATAGTAAGT AGGTCTGTGT GTGTGTGTAG  
7601 GTGCATGGGG AATAAAAAAT AATAAGCAAA TAATGAACAG GGTAAATTTCA  
7651 AAAAGCAGAA AGAGCTATTC AACAAAACCTA CCTGCCTTTT ATTAGATGAA  
7701 ACTCTCAACT CTATGGTTTG TTCTCTCCTG TCAATCTGT TAAATGCTGT  
7751 CAGCTGTTT TCCTTATCAC CCTGGCCACG ACTTCTGTCT TTTCTGCTTG  
7801 GTCCTGTAGA CTCTAACCCA AGGCTCATTC TCTGCCTGGC TATCTGCCTT  
7851 CTGTGGCTCT TTGCCACTAC CTACATTTTC TGTGTTGCAC AGGGAAGGAC  
7901 CATTCCTGT GGACCATAAA ATTCTCTTTT TGAAAGAATT CATTCTTGAT  
7951 TGGGCCACAG CACATCTTGT GAAACAGCAT TAGACATTG CCACTGCTCA  
8001 GCAGCTCTGG GGGAAATGT TTAGTGAGAA GCGTACAGTA GTTTTTTTGA  
8051 CTAACCATGG TGCAACCTCC TCCAGAGGG AAACCTATGA GTATTTCAAG  
8101 GACATGTGAT GGTCTGTTTT TGTCCCCAGT ATCTGACATG ATGGGTAGTG  
8151 TAGAGCAAGA GCTTACAGAT AATGGCTAAA TTAAATTTTC TTTTGAATT  
8201 TTAATATTCA ACTTTTTAGG GTACCCAATC TCCATATTTA GGAAATAAAA  
8251 TTACATAAAA AGTGGAGAGT TTTTATTGTG AAACCTGCACC TCCATATTCC  
8301 CAGTGGTGCA GGATGAGGGA GCACAGGTGT TGGTCTGGGG AAGCCAGGGC  
8351 CCTCTGTGGT TCTGGAGGGT GAGGATTAAG AGGAAGCCTT AGATAGTATT  
8401 TATGAGTATC TGCTGACTTC TCTCTGGGAC CCAAGATCAC TGAACTTTGT  
8451 CCTATTTTGA GATCATCTTT CCAATCCAGC CACTAACAGC TGAAGGATAG  
8501 GCTTGCCCTG AGGCCATTGT AGTGGTTGGA TGAAGATAAA AGATAAAAAA  
8551 CTGTGAGGGG AGGTGTCACA GAAGAAAGGG CCCATGTGGG CAGATTTTCA  
8601 TTCAATTCCT AGTCTTTATT ACAGCAATTC TCCAGTGCTG CAACCTTAGA  
8651 AAAGGATTCC TACAACACAA TGTAGGTACC CATCAGCAGC AGATTGGATA  
8701 AAGAAAATGT GGTACATACA CACCATGGAA TACTATGCAG CCATAAAAAA  
8751 GGAGCAAAAT CATGTCCTTT GCAGCAATAT GAATGCAGCT GGAAGCCAAT  
8801 AACTTAAACG AATTATTGTA GAAACAGAAA AACAAATACT GTGTTCTCAT  
8851 TTACAGGGGG AGCTAAACCT TGGGTAAATG GGGCATAAAG ATGGGAACAA  
8901 TAGACACTAG GGACTCCAAA AGGGGGGAGG GAGGGAGGAG GGCAAGGGCT  
8951 GGAAAGCTTC CTACTGGGTA CTTTGTTTAC AACCTGGGTG ATGGCACGAT  
9001 TAGGAGCTCA AACCCAGTA TCACACAGTA TACCCTTGTA ACAAGCTGAT  
9051 GGTGTAACCC CTGAATCTAC AATAAAATTA TTTTATTTTA AAAAATCATT  
9101 ATAAGGATTT TTA AAAAGAA GGATTCCTAG ACAGGTGCAG CCAACAATT  
9151 TTTTTTAAAT GTTGGCAGGC CGCCACCGCC AGTCACTTAT GCTGCAATAG  
9201 CCCATGTCCC AACATTCCCA ACCTACTTCT CTCCAAAAGA GAAGCTATAC  
9251 TTTCAGATGG CCCTGTGCTG GGTTCCTCCT GGAAGTTTCT GGGGAAAGGG  
9301 GCTTGAGTTG CCCCAGCTG ACTCTCCTG GAGTGGGAGC CGGGGCTTCT  
9351 GATCAGACGT GAGTGAGGCA GGAATCCGC GGTCTCCAG CGCAGCCAG  
9401 AGTGCAGTCC CACGCAGGTC CCGGGTCCTG CGCGCTCGCG CCTTTGCGCT

FIGURE 3, page 3 of 42

9451 GAAGCCGTTA GGATGAGCCC TCTCCTTCCA GAGCTTTAAC CGATGAAGGT  
9501 GCATTGTGTT TGGCGCCCTT GAGGAGGATG CTGTCTTAGG CCTCTTCCCA  
9551 CTGGACGTGT GTGGTGGGCA GAGATCCCGT TCGTCGGTCG CACTTCCACC  
9601 CCGCTGGGGC TCACTCAGGC CGCGGAGCTG CGAGGGAGAC ATCCTCGATG  
9651 GACTCCCTCT ACGGAGATCT CTTTTGGTAC CTGGACTATA ACAAGGATGG  
9701 GACCTTGGAC ATTTTTGAGC TTCAGGAAGG CCTGGAGGAT GTAGGGGCCA  
9751 TTCAATCTCT AGAGGAAGCG AAGGTGGGTC TCACTGGGGC TGTAATCAGA  
9801 GAGACGTTGG GGCTGGGAGC CCTGGAGAGG CATGGGCAG AGAGGGCAAA  
9851 ATTTACATGT TGTCAAGCTT GACCTGGGCC CACTGCAGTG TTCAGGTGGT  
9901 TGACCAGCGT TACCGTTTAT TAAGAATAAC AACACAGCTA ACACATTTCT  
9951 CAAGTATTTT TCTCCGTTTT CTCCTTGGCT GTAGTAAAT CTCCAACCTC  
10001 AGATTGCTCT CAAGATGTTG GCTACATACA GCCTTGCTT AGGAGTCACC  
10051 TTGTTCAATG TGCTCACCTG TCATTAGTCA CCCAGAGGGG CGTCTAGGCT  
10101 AAAGATGCGC CCTCCCCAGT TCAGAGAACT GGAATAATCA CTCTACGTGT  
10151 ATTTGGGAGT GGGGTGGTGA TTGGAAATTT TCTGATGTTA TGTTTTGGTT  
10201 TCTGTTCCCTG GAAGGGGGCA GTGGAAGTGG CTTTTACTCT CGGGTTTCAC  
10251 TAGTGCTGAG GTTTCCTCAT AATATGCCTT AATTGATAGA CCCTAGTTAT  
10301 CAGTACCGAG CTTAGGCTAA CCCTTCTCTT CCCCAGAAGG CTAACCTACA  
10351 GGCTCCTTCT CAGCATGTTG TGCTTCGTAC ATACTCCTAT TGCAGTATTT  
10401 CCAAGTCATT TTTCAATTTG AATTTATTAT TGTATATAAT AATTACTTTA  
10451 TAAGTATATT TGCTCTTTGG ATGTTTGACC CGGTAGACTG GGAGATCATG  
10501 AGCATGTGGA CTATTGAGTT TATTTTGGAT AATTGGTACT TCGTGCCCAA  
10551 AAAACTGTCA GTTGAGTTCT GTCATGTTGA AATTTAGTAA AACTCTTTCT  
10601 ATTAGCCATG TGAACTTTGG GAATATTGAA GCATCCATTC AGTCATGGGT  
10651 CAGTCTAGT TTGAGCACAT TCTATATTCC AAGCCCCATA CCCTGGTATC  
10701 CTCATCTGTT ATATCAGAGG CCTGGACTGT GTACTTTCTG TGGACCAATT  
10751 CAGTCCAAAA TGTATTCTCT GCAAAGCTTA TCTGGATTTT TAATTCCTAG  
10801 AAAAAAGCAG TGTTCCTCTT TTTAAAGTTA AGTGTCTCTG TTCAGGTGCA  
10851 GTGGCTCATG CCTGTAATTC CAGCACTTTG GGAGGCCAAG GCAGGTGGAT  
10901 CACTTGGGGT CAGGAGTTCA AGACCAGCCT GGCCAATATG GTAAAACCCC  
10951 ATCTCTACTA AAAATGCAAA AATTAACCGG GTGTGGTGGT GGGTGTGTGT  
11001 AGTCCCAGGA GGCTGAGGCA GGAGAATCAC TTGAGCCTGG GAGGCAGAGG  
11051 TTGCAGCAAG CTGAGATTGC ATCACTGCAC TCCAACCTGG GTGACAGAGT  
11101 GAGACTCCAT CTCAAAAAGA AAAAAAAAAA GTTAAGTGTT CTTCATATTT  
11151 GTTTAAAGAC ACTCTTATAT TTAGATTTGC AAGTGTAAGT TGTATTGTGT  
11201 TATTTGATAC AAAGTAGCCT TTCATAAGAA ATTCTGGGTT AGCTATCAAG  
11251 TCGAATCTTT TGAACACAT TTCTTCCTTA TTGAAACAAA AGGTTTGTAG  
11301 AGCTGTCTTG CATTTTGGC AAGGACGCTT TGTGTACCTA GTGGTGACTG  
11351 AGGAGGGTTC ACATGTCAAA ACCCAAGGGA GGGGTGTCCC CAGAGAATTC  
11401 TGCACCAACC ACACAGAACA TTCTGTTTCA GAGGAGCACC ATTGTGACTT  
11451 TTCCTCAAGT GGCAGTCACA TCGTTAGGAG GTTTTGATGT GAGGTCTCTT  
11501 CCCACACGTC TCCACCTCCC CAGTAGGAAA ATTTGTTTAT ATAGACAAAA  
11551 CTCAACTGAT TAAAAAATAA AAAAAGAAAT GATACTTACA TTGTCGTGTT  
11601 AAGATACAAA AGCAATAACT TTTTATTGTG AAAATAGTCT GTTTTTGAAC  
11651 AATATATTGT TTTGTTTTTT CCTGTGAAAG TTGAGAAACT AAATATACGA  
11701 AGAGATAATG GTCAGACCAT AAATAAAAAA AGAACTTTGA CTCAAAATTT  
11751 ACAGCAGTCT GCCCAGAAAA CCAGCCCTTT ATCTAAAATA AACAGACCAG  
11801 GAAACCAGCC TGTATGTCA GACTTATAGG AAGTCAGGTT GCTATCTCTA  
11851 GAGACAATAC ACAAAGCTAT GCAATAACTG CTGTAACAGC CCCAAATGGT  
11901 CAGAAATTGA TTAATAACCG ACAGCCCCC TAATTTTTTT CTTCACTNNN  
11951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNTC  
12001 ACCGCTTGCT AGAACTGTGG CCTTGGGTCA TGTTATTTAA TGCCTGGAGG  
12051 CCTCAAATGT TAAGTAGGTA ATGGTAAGAC CTACCCAGTA ACTTAGCATA  
12101 AATAGTAAAT TCATTCAATT AATGTTTTCA AACAGTGCCA GACATTGTTT  
12151 AATGAACTGG GGATATAGTG GTGAACAACA CTGACAGCGT TCTTCATTGT  
12201 ATTCTCAAAA CCCTCCCTAT AGTAAGTAGG TCTGTGTGTG TGTGTAGGTG  
12251 CATGGGGAAT AAAAAATAAT AAGCAAATAA TGAACAATAA AATTATTTTA  
12301 TTTAAAAAAA AAGAAATGAT ACTTACATTG TCGTGTTAAG ATACAAAAGC  
12351 AATAACTTTT TATTGTGAAA ATAGTCTGTT TTTGAACAAT ATATTGTTTT  
12401 GTTTTTTCCT GTGAAAGTTG AGAACTAAA TATACGAAGA GATAATGGTC  
12451 AGACCATAAA TAAAAATAGA ACTTTGACTC AAAATTTACA GCAGTCTGCC  
12501 CAGAAAACCA GCCCTTTATC TAAAAATAAC AGACCAGGAA ACCAGCCTGT  
12551 TATGTCAGAC TTATAGGAAG TCAGGTGCT ATCTCTAGAG ACAATACACA

FIGURE 3, page 4 of 42

12601 AAGCTATGCA ATAAGTCTG TAACAGCCCC AAATGGTCAG AATTTGATTA  
12651 ATAACCGACA GCCCCCTAA TTTTCTCTT CACTTCCAAC TTAGGACGAA  
12701 CCAGAGAAAG CTAAATATGC ACCACCTACT AATCAAATAG GGTGCCGCGT  
12751 TTCTAATGAA CCCTCCTACA GCTTCCCCAG GCCAGCAGCC CCAATCAGG  
12801 AAACGCCCTGA AGCCTTCCCT TTTTCTCACT GTAAAGCTTT CCCACTCCTC  
12851 TGCCTGGCTT TGAGTCTCTG TCAATACACA AGTGAGGGTG TCTGACTCCC  
12901 TTGCTATAGC AAACCTGGGC CAAGTAGATT TTACTTTTCT CATTTGATTG  
12951 GTCTTTTATT TCTAGAAGGA ACATACAAGA AAATTTAAAG GGAATCCAT  
13001 TCCTAATCTT TCAATATATA GTAGTCCCT TTTATCTGCA GGGCATATTT  
13051 TCCAAGACCC CCACTGAATA CCTGAAACTG TGGGTAATAT TGAACCTAT  
13101 ATATACTCTC TCTATATATA CATATATATA TATATTTTTT AATTTTTTTT  
13151 TACTTTTATCT TTAATTAGCT TTAGCTCTTT TTTTTTTTTT TGAGATGGAG  
13201 TCTCACTCTG TCACCCAGGC TGAGTGCAGG GGTGCAGTCT TGGTTCCTG  
13251 CAACCTCTGT CTACCGGGTT CAAGCAATTT CTTGTGCCTC AACCTCCGGA  
13301 GTAGCTGGGA CTACAGGCGT GTGCCACCAC TTCTGGCTA ATTGTTTTAA  
13351 ATTTTAGTAG AAACGGGATT TCACCAAGTT GGCCAGACTG GTCTCGTACT  
13401 TCTGACCTCA AGTGATCCGC CCACCTTGGC CTCCCAAAT GCTGGGATTA  
13451 CAGGCGTGAG CCACCATGCG CCCAGCCATA GACTATATAT TTTGATCTG  
13501 ATAAGTGGT CAGCTACTAA GTGACTAACA GGCAAGTAGC ATCTATAGTG  
13551 TGGATATGCT GGACAAAAGG ACATTCACCT CCTGGGCAGG ATGGCACAGA  
13601 ATGTTGAGAG ATTTTATCAT GCTACTCAGA ATGGTGTGCA ATTTAAACT  
13651 TATGAGTTGT TTGTTTCTGG AGTTTTCCAT TTAATAGTTC AGACCATGGA  
13701 TTGACCGCAG GTAAGTAAA CTGTGGAGAG TGAAACTGTG GATAAGGGAG  
13751 GACTATTGTA TTGTTAAGTC AGACTCATT GGCAATCATA ACTCTTGATT  
13801 TGCCATCAGA AATGCTGCAG AAATATGGGT TAAAAAAAAC TGTCAAAAA  
13851 TAGGGTCAGG GATGTCCTTT AACTTGTTAC TTCCAAAATG TTAGTGAAAA  
13901 CTGTGGCCCC AAAGAGTGAA AGGAACAAAT GACTAAGAGA AAATCTTGTT  
13951 TTCAGGATGA CAGATTAAAA AAGAAGCAAC TTGCTGAAAC ACTGAAAATC  
14001 TCTCCACTTG TAAGATAACA CAAACTGGC TAAACTGGT TGGAAATGAAT  
14051 ATGGCCAACT CAAGTCTGCA CAGAACTAAC TTGGTGATGT TACAGCCCAA  
14101 ATTTCCACCA CATATTTTAT ACTAACTCCC CCCGGATTTT CACACATGAT  
14151 CTGTGAGGTA GCATGAAGAG GTAAGTATGC ATGCCTAAGG ACTTGGGAGA  
14201 CCTCCCCATT TCCCTGACC AATCACCAC TAATCCCAGA ATCCGCCCCC  
14251 AAACCTTTTC TAATAACTAC CTTAAAGCCA GCATAGGGAG ACAGATTGTA  
14301 GCTGGACTCC TGTCTTCTTG TGGGTACCT TGCAATAAAA AGCTTTTCTT  
14351 TTCTCAACAC CTGGTATTAT AGTATTGACT TCTAGTTCAT CGGGCAGCAA  
14401 GCCCCTTTTG GTCGGTGACT ATTCTTGTTT GCTGATATTT CCATTGGCCA  
14451 AAATATAAAC CTCTTAGATG AAACCTCAGT ACGTAAATGG CGCCACAGAA  
14501 TGCTGTGACA TTTTCTCTT GGATTATAGC AGGTTACTTT ACTGAATACC  
14551 GTAGGCAGTT ATAACACACT AAGTATTTGT GTATCTAAAC ATAGAAAAGA  
14601 TACAGTAAAA ATATGGTAAT TTTTTC AAC TTTTAGTTGA GATTGAGG  
14651 GTATGTGCAC ATTTGTTACA AGGTATATTT GCATGATGCT GAGGTTGGG  
14701 GTACAATTGA ACCCTGTCAC CCAGGTAGTG AGCATAGTAC CCAATCGATA  
14751 ATTTTTCAC CTTGTCCAT TCCCTCCCCG TTCTTGATG CCCCAGTTT  
14801 TGCTTTTCCC ATCTTTATAT CCGTGTGCAC CCCATGTTTT GCTCCCATGT  
14851 GTATGTGAGA ACTTGTTGTT TTTGGTTTTT TATTTCTGCG TTGATTCGCT  
14901 TAGGATAATG GCCTTCAGCT GCATCCATGT TGCTGCAGAG GACGTGATTT  
14951 TATCTTCTT TATGGCTGTG TAGTATTCCA TGGTGAAAAA TATAGTACTA  
15001 TAACCTTACT AAATCACTGT CATATATATG GTCTATCATT GACTGAAATG  
15051 TATACAGTGC ATGATATATA TATATATATA TCTATAATGT CTTATCCATT  
15101 TCGTGTATTA TGAGATTTGA TTGCTAATAT TTTATACAGG AGTTTTGCAT  
15151 CTTTTTCACT AGTTGACATT GCTTGTAATT TTCTTTTTTT TGTGATGTC  
15201 CTGTTAGGTT TTAGAATCAA GTGTATACCC GCCTCATAAA ATGGGTTGGA  
15251 AAATGTTCCC ACCCTTTCTG TTCTCTGGAA AATGGGTGTT TTTTCTTAA  
15301 AGTTTGGTAG ACATTATGT TAAAACCATG GGGTCCCTCGA TTTTCTTCA  
15351 TGGAAATGTT TTCAAATTAC ACTTTAAAT TCTTTAAAT CTGAGTATAG  
15401 GGCTATCAGA CTTTCTGCTG TCTTATGTCA GTTTTAAATA AGTTGTTTTT  
15451 GTAGGCGTTT TTTATCTCAC TTTTATATTT TTGATATAAA GCTTTTCATA  
15501 ATATCATTAA GTTCTATAGT GTCTAGTAGT TTCCATCTTT ACTTTCTGAC  
15551 ATTGGTTATT TGCCAGTTT AGGAGTTTAT CAATTTTATT AGTCTTTTCA  
15601 AAGAACCATC TTTTGGCTTT GTTAATCCTC CCAATGGTGT GTTTTCTTTC  
15651 TCATTACTTT TTGCTCTTTA TTTCTTCAA CTTCTTTTTT GCTTAATTTT  
15701 AAAATAATTT CTTGAGATTG AGATAAGCCT CAATGATGGG TCACCGATTT

FIGURE 3, page 5 of 42

15751 CCAGTCTTTC TTCTTTTCTA ATTATGCATT TTAAACCAGA AATCTTTCTC  
15801 TAAGTGTAGC TTTAGTTGCA GCTCACAAGT TTCAGATCTG TCTCTCAGTC  
15851 TGGAGGTTGG AGATCTGACC ATGACCATGA AACCATCCAG TCACAATGTG  
15901 GCATTATTTT TTTAATTTTT TTTTTTTTTT TTGAGATAGA GTTTCACCTC  
15951 TATTGCCCTAG GCTGGTGTGC AATGGTGCGA TCTCGGCTCA CAGCAACCTC  
16001 CACCTCCCAG GTTCAAGCGA TTCTTTTGCC TCAGCCTCCC AAGTAGCTGG  
16051 GATTACAGGC ATGCCCCACC ATGCCCAACT AATTTTGTAT TTTTAGTAGA  
16101 GATGGGGGTT CTCCATGTTG GTCAGGTTGG TCTTGAAGTC CCGACCTCAG  
16151 GTGATCCGCC CACCTCAGCC TCCCAAAGTG CTGGGATTAT AGGAATGAGC  
16201 CACTGTGCCC GGCCCAACTT GGCATTATTT ACCCAGAAGA GCATGACCAT  
16251 GAGAACAGTA GAATTTGTAA GCTTTGAGTG GGTGACTATG AGTGTCATAA  
16301 TAGGTAGATA GGTATATATT TGGGTGGTGG TAGGAGAGGG CTTACAGTTT  
16351 GCTATGACAG CTTTTTATAT GGATCATCCT TAGTAAAAGA TTATTTAATT  
16401 TTTGAAATCA AAGGGGAAAA CACTAGTTTA GGCTTTCTTC TTTCTTTCTT  
16451 TTTTAGAGAC AGGGTCTTGC TCTGTCACCA GGTTAGAATG CAGTGGTGCA  
16501 ATATTGCTCA CTGTAACCTC AAATTCCTGG GCTCAAGTGA TCCTCCTACC  
16551 TCAGCCTCCA AGTAGCTAGT ATTTACAGGC ATGCACCAAC ACATCTGGCT  
16601 AATTTTAAAA ATTTTTTATG GAGATGAGGT CTCACTATGT TGTCAGTCT  
16651 GGTCTTGAAT CCTGACCTCA AGTGATCCTC CCCCATCAGC CTCCCAAAGT  
16701 GCTGCAATAT TTTAAATCCT GTGGTAGGTC AAGTGGTTGT CTTCTATCTT  
16751 GGGGTTTATA AAGTACATGT CAAGAAATTT AGGGTATGGT TAGATTAGCT  
16801 TTAAAAATGT CATGTTTTAT AAAAATCAAT GCATCATTTT TCTGATTGAA  
16851 AATTTAACAC AAGACTCAGA ATCTTTTTGC AGTAGTGGAA TTACTTTTAT  
16901 TATAGATCTT TGCGATAATG AATGATGATA CATCTGGCCA AAAATAGGTA  
16951 CTATAGTCTT TTAGGAAAAC AGCTAATCTG CTTGAAATAT GTGTAGAAAT  
17001 AATTTAGTGC ATCAGCCCAT ATTGGCAATA ACTTCTCTCT AATTTTTTTT  
17051 TATAGAAAAT TTTTACTACT GGAGATGTCA ACAAAGATGG GAAGCTGGAT  
17101 TTTGAAGAAT TTATGAAGTA CCTTAAAGAC CATGAGAAGA AAATGAAATT  
17151 GGCATTTAAG AGTTTAGACA AAAATAATGA TGGTGTGTCT TTCTTTTGTA  
17201 TTTATCACCA GCTATGAAGA AGCATTATATC ATGCTTTCAA GAGTCTAAAA  
17251 GCATGCTTAT TTAATCTCTC TGGTTTTAGA TGATAATTAT TATTTGTGTT  
17301 AATACTTTTT TTTAGTAATG TGATTTTTAT GTAGAGTTA TATTATTAG  
17351 TGAAGAAAAC TTATAGATAG CTTTTCTTTT TCATTACTTT GAAATGTAAT  
17401 GAATTACATT TCTGAATTAA AAAGTGTGGG CAGGGCCTGT TGTAATGTT  
17451 AACTATGGAA CATTATGCTG ATTTGAGTTA AACCTGTAGG TTAAAAATAA  
17501 TAATTATATT TTCTTGCTCT CTGGGTAAAA TGAGATTCTT TTTTATTTGT  
17551 ATAGAAGAA GACAGTTGTG TCATCTAAAA TTTAAAAAAC TTTCAGATTA  
17601 TCTTGCACTT GTTAGTTTTT TTGGAAGAAT TAATTTAGAG AAGATATCTC  
17651 TGATCCTGGA AATTAGGGAA AAATAGCATA TAAACGTTTA AGTGTGTACC  
17701 TTCTGGTTAA GATTATGACT TCTATATTTT GATTAATAGG TTGGAGTTTG  
17751 TCTTAATCTG TTTTCTGTTG CTGTAATGGA GTACCACAGA CTGGGTAATT  
17801 TATGAAGAAA TGAAATTTAT TTCTTATAGT TCTGGAGGCT GGGAGGTTCA  
17851 AAGTTGAGCC GAATCTGGTG AGGGCCTCTT ACTATGTCTT AACATGCTAG  
17901 CAGCATCAC AGAGCAATG CACTACCTCA GATCTCTCTT CCTCTCTTA  
17951 AAAAGCCACT AGTCCCATCA TGGGGGCCCT ACTCTGAAGA CCTTATCTAA  
18001 TTCTAATTGG AAATAGGGTC TTGAAGCCCT CATCACTAGA GGTAACCTTT  
18051 AACAGGAAGA GAGAATTTAT AAAAATTATA ATGCAGCACC AAATCCTCC  
18101 CTAATTGTGA ATAGTCAAGG TCATTTTATT TACAGACTTG TTATTAAAGA  
18151 AACAGGTTAA ACAAATAGAT TGAGAGGAAA TGTGGTTTCT GTCTGAGATC  
18201 AGCAAACTTT TTTGTCCAGA AGTCCAGATA ATAAATATTT TAGCTTTGTG  
18251 GGTCAATGTT TCTCAGTTGT AGCTACTTGT CTCTGCTGCT GTACCTCAAA  
18301 AGCAGCCATG GATAATATGT AAATGAATGG GGATGACTGA TTTCCAATAA  
18351 AAATTTTATT TACAAAGATA GTTAATACAC CTTATTTGGC TTGAGGGTTA  
18401 TAGTTTGCCA TCCCCTGATT TACAATGAAT ATTAAAGTTT AATTCAAAGC  
18451 AAGTTCCTTC AAACAAACAA ACTAACTCT AGATGATTTT GAAGATTATT  
18501 CACATCTGTG ACTCTCAGCC AGGAAGAGCT GAGTTTGGGT TGGAAAGTAG  
18551 TACTATTGGA ACATTTGTTG CCCATAAGCC TTACAATATA TGCCCTAAG  
18601 TCTAGCCTTA GTCCAGTCTT CTAGCAAAAC TCAGTTTTCT TTCTTCTCTG  
18651 CAAACTTTCA TTCCACATC GACCTCTGCG AGTTCAGATT GTCTGTCAGG  
18701 TCAGATTGTC TGTGTGCTGC TATGGTAGGC AGTAGCTGAG AGATGGAGCT  
18751 ACCTTAAGAT CAATTGCCAG ATAATCAGAG GTCAATTATC CCAGTGACATA  
18801 AGTAGTGTAC ATATCAATTG TTCAATTTAT AAAATTCTAA ATGAACCAGA  
18851 GGCAATAATT AAAGATGAAA TTTTGATGGT ATATTTGTAG GAAATCTACA

FIGURE 3, page 6 of 42

18901 CAATGTTTCC CTAATTTCCC ATGTTTGTGT ATTTTAAAAC AATGTGGCAT  
18951 TATTGGTTCA TATTTTATT TTTTAGACTT CCTTAATGCA AAACATATAC  
19001 AGTTGATCCT CATTATTTGG GGATTCTGTA TTTGCAAATT TGCCTACTCA  
19051 ATAAAATTTA TCCCCAAGT AACCCCAAAA TATATACTCA CAGTACTTTC  
19101 CCAGGCATTC ATGGACATGC ACAGAGCAGT GAAAAACTTG AGTTGCTCAG  
19151 CATGTACATT CCTAGCTAGT AGAATAAGGC AATACTCTGC CTTCTTGTTT  
19201 CAGCTCTCAT ACTATTAAGT AGCAAGTATC CCTTTCAAGG TCTATTTTGT  
19251 GCCAGTTTTT GCATTTTGTG ATTTTGTGTG GTAATTTTCT TTTTAAAATG  
19301 TTCCCCAAAG GTAGTGCTGA AGTGCTGTCT AGTGTTCTTA AGTGCAAGAA  
19351 AGCCATAGCA TGCCTTATGG AGAAAATATA TGCCTTGGAT AAGCTTTGCC  
19401 CCAAATTCAG TGTTAGTGAA TCAACAGCAC ACATTAAATG AGGTGCCCTC  
19451 AAACAGAAAC AGACATAAGA CATGGTTATG TATTAATCAG TTGATGAAAG  
19501 TGTGTAAATC AGAGGCTCAC AGGAACCTAA CCCTGTTTTT CCTGTAGGAA  
19551 CAATGGTTTG GTATTTGCTA ATTCAGTGTT TGCAATGAAT ATAGAACTTT  
19601 ATGGAAGATG ATTGCTGTGA ATAATGAGAA TTAACCATAT CTCTTTAAGA  
19651 GTGCATTTCT AAAGGAGAAT ATTCAGAAGG GTATTTGCAT AATTTCTTTA  
19701 CTAACAGATG CTGCCTCTCA CTGTCCTTAC ATGGTCCAGA TTCTCATGCT  
19751 GTCCTTCCCC TCTCCCCAGG AGGATTCTCT CAGAATCCTG TCATCTCCTC  
19801 CAGGGTCTCT TCTCCAAGAA AGTCTATCCT TTCACCACTA ACAGTAATTT  
19851 TGGTCTTCCCT CTTTTTCTGG AGAAGTCAGC TGTTTATGCT GCTTCAGCAC  
19901 CAGACCCTCT CTTACTTTGT TTTGTTTCAT TCTTTTTCAT GTACAGTAGT  
19951 CTTAGGATTC TCATGAGCCT GTGAGCTGCT AGAAGGAAAT ACAGCAGTGC  
20001 TTACATTTAT TGCTTCTATT TTATTTTCTA TTTTCTCTTC CTGTCTCTG  
20051 ATTGTTCTCC TTCTGTCCAC AAACATGCTC TAATTTCCCT AGTATTAATA  
20101 ATTTTCTGTC TTTTGTGTG TTTTATCCT TGCTCCCTTA TTTTACTGCT  
20151 CAGATTTTTA TTTTATTTA TTTATTTTG AGATGGAGTC TCACTCTGTC  
20201 ACCCAGGCTG GGGTGCAGTG GCGCGATCTC AGCTCACTGC AACCTCCGCC  
20251 TCCCAGCTTC AAGCAATTTT CCTCTTTTAG CCTCCCAAGT AGCTGGGATT  
20301 ATGGGCACCT GCCACCATGC CTGGCTGATT TTTCTATTTT TAGTAGAGAC  
20351 GGGGTTTCAC CATGTTGGCC ACACCTGCTCT CTAACCTGCTG ACCTCAGGTG  
20401 AACCAACCCG CTCAGCCTCC AAAAGTGCTG GGATTGCAGG TGTGAGTCAC  
20451 TGTGCCTGGC CTTTACTGCG CAGATTTTTA AAAGAATAGT CTGTGCTTTA  
20501 GCTCTATTTT CTTACTTACT ACTTCTCTTT AACTCAGTCA TATATGATGT  
20551 TTTGCATAGT AAATGTCTAG TAATTTATTA AAAATGTAGA AATAGGTACT  
20601 TTTAAATGA ATAGATCCTA CTTTAATTGA ATTTATCTTG GAGTTAGAAT  
20651 ATCTTGATTT GGATTTTAGT TCTGCTACTT CTTAATTACA TTACTTGGTA  
20701 AGGCCACTTG TGAAGTCAGT CTCTTTGGAG GAATATTATT TATCTATAAG  
20751 GCTGTTACAA TTACTGAATT TAAAAAATG TGTATTTATT TTTTAAATGA  
20801 TTTGTTACAT TTTTAGTATT GATGTTGGGA TAGGCATTTA AGCAAGTCTA  
20851 TAACTCACCT ACATGCATAA TTTTGCCTTA ATCAGTTTAA AGCTTTCTCT  
20901 TAAATGAGAG ATTTGAAATT CATAATTTCT GTGGTTCTTA TCAGTTCTGA  
20951 GTTTTATTTT TTGCCCTTTT TATTTTTTTA AAGGAAAAAT TGAGGCTTCA  
21001 GAAATGTGCC AGTCTCTCCA GACACTGGGT CTGACTATTT CTGAACAACA  
21051 AGCAGAGTTG ATCTTCAAAA GGTAAGCTCT TCATGTTGGT CAACAATTGA  
21101 CTTTCACTTT AATATCCTGC ATTAGAATCT TGTGTTTGTA AGTGTGGCTT  
21151 TAAAACACCT CCCTAGTCTT CATTATGTAT ATCCAAGATC TTTTGTCTT  
21201 TTTTCTCTCC ATTCATTTTG TATGTGTACA TTTATCTAAA GTGTAAGAAT  
21251 GGGAAAGTGA AGCTCAGACT GGACTCTTTC TTTCAAGGCC TCAAAGGATA  
21301 GTGGAATGGC AGGAAGTAAG GTTTAACTC CATAGATGAG GAGCTGAAGA  
21351 GTTTTGGTGT TGCTTTTCTT CCATTTGATT TCTAATGTGA CAGTAAACT  
21401 CATTGATTCA AACTAAGAAG ACTAGCAGAT TCATCACATT ATTTAACCTA  
21451 GATGTGACTG GAAAAAAGGG AAATTACTAA GCTCTCCAAG CTAACAAAGA  
21501 AATACCTGTT TAACTTTCA GAAAACAGAA ATGCAATTTT GAACCTTATT  
21551 GTCTGGGGCA ATCAGTTTGA CTATTTAAGT CAGACTTTTA TACTCTTAAT  
21601 GTTTTGTTC ATGGGATAGA GCAGTAATCT CTGCAGCCCA GGTGCTCTCA  
21651 AATACTCTGT TGCTATAAAC ACAGGCGAGG AACTGATTTT TTATGATAAC  
21701 GTAAAACAGA AAAGGACAAT TATATTGTAT TAATATTGTT GTGAATATTT  
21751 TCAGTCTCTA CATTGTCTAA AAATCTTCT AAATGGCTTT GTTATTGAAT  
21801 TTATCTCAT TTTATCTGT GCCAACAGCA TTTTCATCCT TTCTCTTCAT  
21851 AATTTCTTTT ACAAACAGCT GCTCAAGAGG AAGGCTCAAA GTCTCAAGGC  
21901 TGAGCACGTA ATGACTTTTG TTAGTACTAG ATGAGAAGGG CTTTCTGAG  
21951 GAAATGAAAA CCTAAAACAT GAAAAGAAGA TAAACAGAAT TTGGACAGTG  
22001 AGATATAGAG CATATAATAT TCTGCTTCTA AAGTAATATT CTTCTAGGAA

FIGURE 3, page 7 of 42

22051 AGTGAGGGCG TTTCCCTGGC TGTAGGCCA GAAATCATAT TCCTATATTT  
22101 TCTTTGATAG CTTTAGGAAT AATGCAAATT CTAAGCCCAA GCTTCAGAAAT  
22151 AGACTAAGAA GTATTAGCTT AGCTGCCATG ACAAATACC ATAGGCTGGA  
22201 TGCATTAAAC AATGGAAATT TAGTTTTCA CAGGTCTGGG AGCTGGGAAG  
22251 TTAAAGATGA GAGTGCCAGC ATGGTTGGGT TGTAGTGAGG GCTCTCTTTC  
22301 TGGCTTGCCAG ATAGACCCCT TCTCACTGTA TTGTCATATG GCAGAGAGAG  
22351 AGAGAGAGAG AGAGAGAGAG AGAGAGAGGG GATCTTTCTC TTGCTTTCTA  
22401 TTATAAGGCC ATAGTCCTGT TGGATCAGGG TTCCATTCTT ATGACTTTAT  
22451 TTGACTTTAC CCCCTAAGA TGCTATCTCC AGATATAATC ACACGGTGGG  
22501 TTAGGGCCTC AACATTTGGA TTTGGGAGGG ACACAGCTCA GTCCATAGCA  
22551 AAGGATAATG CAGAGGGTTG GATATTTAAA AGTAGCTACA CAATTTTTAA  
22601 TATAAATATT TTATGGTAAC TTTTTTTTTT TTTTGAGATG GAGTCTAGCT  
22651 CTGTTGCCCA GGCTGGAGCG CAATGGTGCG ATCTCAGCTC ACTGCAACCT  
22701 CCGCCTCCCA GGTTCAGCA ATTCTCCTGC CTCAGCCTCC TGAGTAGTTG  
22751 GGACTATAGG CACGCGCCAC CACGCCTGGC TATTTTTTTT TTATTTTTAC  
22801 TAGAGACGGG TTTGCACCAT ATTGGTCAGG CTTGTCTCGA ACTCCTGACA  
22851 TCAGGTGATC CACCCATCTT GGCCTCCCAA AGTGCTGGGA TTACAGAAGT  
22901 GAGCCACCGC GCCTAGCCAG CAGCTTTACT GAGATGTAAT TCACATGCCA  
22951 TAAATTCATT TTTCTAAAGT ATACAATTCA GTGACTTAAA ACATTTATTT  
23001 ATTTTTAAAT TGACAGAATT ACATGTATTT ATCATGTACA ACATGATGTT  
23051 TTGAAGTATA TGTACATTGT GGAGTACTA AGTCTAGCTA ATTAACATGA  
23101 TACATCTCAT ACTTAATGAT TTCTGTGGTG AGAACACTTT ACATCCATTG  
23151 TCTTAGTATT TTTCAAGAAT ATAATATATT ATTATTAATT GTAGTCTTCA  
23201 TGTGTATAG TGGAGCTCTT GAACTTATTC CTCATGTCAA GCTGAAATTG  
23251 TGTGCTCTTT AACACAAACC ATACCCGACT CCCAAAGTAT TCTGCTCTCT  
23301 GCTTCTATGA GATTAACCTT TTCTGATTCC ACATGAGTGA GATCATGCAG  
23351 TATTTATTTG TCTTTACCTG GCTTATTTCA TTCATATTGT TACAGATAAC  
23401 AGGATTTCTT TCTTTTTTTA ATGGCCGAAT AGTTTTCTAT TGTATATGTA  
23451 TAGCACATTT TCTCTCTTCA TGCATTGGTG GACACTTAGG TTGATTCCGT  
23501 ATCTTGGCTA TCGTGAATAG TGCTATAATG AACATGGGAA TGCACATGGC  
23551 TCTTTGACAT ATTGATTTCA TTTTATATAT GTGTATATAT ATATGTATAC  
23601 ACACACATAC ATACAGTGGT GGGATTGCAG GATCATATGG TAGTCTATA  
23651 TTTAATTTTT AAAGGAACTC CATACTGCTT TCCATAATGG CTGTATTAGT  
23701 TTAACCTCTC ACCAACAGGG TGCAAAAGTT CCCTTTTCTC TACATACTTG  
23751 CCAACACTTG TTATCTTTTG TCTCTTTGGT AATAGTCATT CTAAGTGTAG  
23801 TATGAGGTGA TATCTCATTG TGGCTTTTAT TTGCATTTCT GTGGTAATTA  
23851 GTGATATCGA GCTTTTTTTT TTTTTGTAC TTTGGCCATT TGTATGCTT  
23901 TGAAAAATGT CTATTGGGGT TTTTGGTTG TTTATTGAG GTTTTNNNNN  
23951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN  
24001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN  
24051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN  
24101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN  
24151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN  
24201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN  
24251 NNNNNNNCCG GGGTCCCGT CATCTCCCT GCCTCAGCCT CCCGAAGTA  
24301 GCTGGGACTA CCAGGCCACC CGCCACCAC GCGCCGGGCT AATTTTTTGT  
24351 ATGTTGAGTA GAGACGGGGT TTTACTGTGT TAGCCAGGAT GGTCTTGATC  
24401 TCCTGGCCTC GTGATCTGCC CGCTCGGCC TCCAGAGTG CTAGGATTAC  
24451 AGGCGTGAGC CACCGCGCCT GGCCTGATTT CTAGTTTTTT ATTATTGTGG  
24501 TCGGAAAAGA AACTTGATAT GATTTCAATC TGCTTAAATT TGTTAAGACT  
24551 TGTTTTGTGG CCTAACATAT GATATCCCCT GGTGCATGTT CCATGTGCAG  
24601 TTGAGAAGAA TGTGTATTCT CTTGCCATTA GGTGAAATGT TTTATGTCTG  
24651 ATCTGTCCAT TTGTTCTAGA GTATAGTTTA AGTCTGATGT TTCTTACTGA  
24701 TTTTCTGTTG AGATGATTTG TCTATTGCTG AAGGTAGGGT GTTGAAGTCC  
24751 CCTACTATTG CTGTATTGCA GTCTCTCTCT CCTTTCAGAC GTATTAATGG  
24801 TTTTTATTTT ATTTTATTTG TTGTTGTTGT TGTGTTGTT GTTGTTTTTG  
24851 AGACGGAGTC TCACTCTGTC ACCAGGCTGG AGTGCAGTGG CAGGGTCTCG  
24901 GCTCACTGCA GCGCCGTCT CACGGTTCAA GCGATTCTCC TGCCTCAGCC  
24951 TCCCGAGTCG CTGGGACTAC AGGCGCATGC CACCACGCCC AGCTAATTTT  
25001 TGTATTTTTA GTAAAGACGG GGTTCACCA TGTGGCCAG GATGGTCTTG  
25051 ATCTCTTGAC TTCATGATCC ACCCGCCTTG GCCTCCCAA GTGCTGGGAT  
25101 TACAGGTGTG AGCCACCACC CCTGGCCAAT GTTTGGTATT TATCTTTAGG  
25151 TGCTCTGATG TTGGGTTTCAT ATATATTTAT AAAAAACAAT AGCTACATAA

FIGURE 3, page 8 of 42

25201 CTTATTAAGG GATATGCAAT ATAAAATATA TAAATTGTGA CACTGAAAAT  
25251 TTAAAATGGG AGGAGTGGAG TAAAAGTACC TTCATATAAC TTACTATTAT  
25301 ATCCTCTTAT TGAATTGACC CTTTATCAT TATATAGGAA CTTTGTCTCT  
25351 CCTTTACAAC TTCTGACTTA AAGTTTGTCT TATATGATAT AAGTAAAGTT  
25401 ACTCCTGCTC TCCTTTGGTT TCTGTTTCCA TGGAAATATCT TTTTCCATTC  
25451 CTTCAACATC AGTCTGTGTG TATTTTACAT GATGAAATGA GTCTGTCATG  
25501 GGCAGCATAT AGTTGGATCT AGTTTTTTTA ATCCACTCAG AACTGTGTGT  
25551 TTTTGATTGG ATAATTTAAT CCATTCATGT TCAAGGTAAT TATTGATAAG  
25601 TAAGGACTTT GTACTACCAT TTTGCTTATT GTTTCATGGT TCTTTTATAG  
25651 ATCCTTTATT CTTTCTTCC TCTCTTGCTG TCTTTTCTTT GTGGTTAAGT  
25701 GATTTTCTCT AGTGGTATGT TTTGATTCTT TGCTTTTAT TTTTGTGTGA  
25751 TCTCCTATTG GTTTTGGTT TGTGGTTACC AAGAGGTTAC AAAAAACATC  
25801 TTAAGAGTTA TAATAGTTTA TTTTAACTTG ATAACCTAAT TTTTATTGCA  
25851 AAAACCCCC AAAACAAAA AATCTACACT TTTACTTAAT CCCCTGAAAT  
25901 TTTGAATTTT TGATGTCACA GTTTACCTCT TTTTCATATT TGTATCCCTT  
25951 AAATTATTGT AGCTATTATT ACTTTTAATA GTTTTCTCTT TCCTACTACA  
26001 GATGTAAGTG ATTTGCATAC CATCATACA GTATTATTTT GAATTTACCT  
26051 GTGTACTTTT TTTTATCAGC CAGTTTATA CTTTCAGATG TTTTGTGTGT  
26101 ACTCATTAGC ATCTTTTCTT TTCAGCTTGA GGAGCTCCTT TTACGTTTCT  
26151 TATAAAATAG GTGCGGTCAT GATTATCTCC CTCAGCTATT GTTGTCTGG  
26201 GAAAGTATCT CTCCTTCATT TCTGAAGGAC ACTTTGCTGG GTACATTACC  
26251 CTTGGTTGGT ATTTTCTCC TTGAACGCTT TAAATATATC ATCCCTTCT  
26301 CTCTGACCT GTTAGGCTCTC TGCTGACCAG TCTGTTTCCA ACCATATTGG  
26351 GACTGTCTTA TATGTTATTT GCTTCTTATC TTTTGTCTGT TTCAGGATCC  
26401 TCTCATTGTC TTTGATTTT GATAGTTTGA TTGTAATATG TCTTGGGGTA  
26451 GTCTTGTGTT GATTGAATCT GATTAGAGAC CTTGGACTTT TCCTGCATGT  
26501 AGATATTTAC CTCTTCTCC AGGTTTGGAA AATTTTCTGT TACTGTTTCT  
26551 TTAATTAAGC TTTTACCCC TTTTATCTTC CTTTCTCCTT TCTTCAACTC  
26601 CTGTGACTCA AAACCTTGCT CTTTGTATGC TGTTCATAA ATCTGTAAAG  
26651 CTTTCTTCAT TCATTTTCAT TCTTTTCTCT CCTCTGTGTA TTTTCAAATA  
26701 ACCTGTCTTT GAGTTCATAG TTTCTTCTT CTTCTTGATC ACTTCTGCAG  
26751 TTGATGCTCC CATATTGCAT TTAATTTTGT TTCATTGTAT TTTTCAGCCC  
26801 CATGATTTCT GTTTGATTTT TTCTTTTATT ATTTTCATCTC TTTATTACCT  
26851 TTCTCTTTGT GGTCACTCGT TATTTTCCTA ATTTTCATTGA ATTGTTTCTT  
26901 TGTATTTTCT TGAAGTTTGC TGAGCTTCTT TTGAATTCTA TGTCAGTTCA  
26951 TACATCTCTG TTTCTTTAGG GATGGTCGCT GGTACTTTAT TTTGTTTCTT  
27001 TAGTGGTGTC ATTTGTTCTT GATTGTTGTT GATGTTTGTG GCCTTGTGTT  
27051 TACATCTGTG CATTTGAAGA AGTAGGCACT TATTTAGTC TTTGCAGACT  
27101 GGCTTTGTCT GAGAATGCCC TTCAACAGTC AGCCTGTCTA GAGATTCTTT  
27151 AATATTTAAT TAAATATCTT TAATATTTTG AAGAACTTCC AAATTGTTTC  
27201 TAAAGTGGCT GCACCATTTT ATAATCCCAG CAGCAATGAA TGAAGGTTTC  
27251 AGTTTCTCCA TAGCTATATG AATACTCATT ACTGTCTGTC TTTTCATTTT  
27301 TTGATTTTTA TTTTTTTTTT GAGAAAGGGT CTTGCTCTGT CATCCCATCT  
27351 GGAGTGCAAT GGCACAATCA TGGCTCATTG CAGCCTCAAC TTCCCTGGCT  
27401 CAATTGATCC TCTCACCTCC TGAGTACCTG GGACTACAGG CATTGTACCA  
27451 CAATGCCTGG CTAATTTTTA TATTTTGTG AGAGATGTGG TTTTGCCATG  
27501 TTGCCTGGTG TATTAGTCCA TTCTCATGCT GCTATAAAGA ACTGCCTGAG  
27551 ACTGGGTAAT TTATAAAGGA AAGAGGTTTA ATTGACTCAC TTTTGTCTGG  
27601 CTGAGGAGCC CTCAGGAAAC TTACAATCAT GGTGGAAGGG GAAGCAAACA  
27651 CGTCCTTCTT CACATGATGG CAGGAAGAGC AGTGCCTAGC AAAGAGGGAA  
27701 AAAAACCTTT ATAAAATAAT CAGATCTCAT GAGAAGTTAC TCACTATCAT  
27751 GAGAACATCA GAATGAGGGT AGCCTCCTCC ATGATTCAAT TACCTCCAC  
27801 TGGGTCCCTC ACGTGACATG TGGGGATTAT TGGAACATA ATTCAAATG  
27851 AGATTTGGGT GAGGACACAG CCAAACCATA TCATTTTTCG CCTGGTCCCT  
27901 CCCAAATCCC ATGTTCTCAC ATTGCAAAAC ACAATAATGC CTTTCCAGCA  
27951 GTCCCCAGC GTCTTAACTC ATTCCAGCGT TAACCTAAAA GTCCAAGGTT  
28001 TCATCAGAGA CAAGGCAAGT CCCTTCTGCC TATAAGCCTG TAAAATCAAA  
28051 AGCAAGGTAG TTATTATACT TCCTAGATAC AATGAGGGTA CAGGCATTGA  
28101 TTAATATATC TTGTTCCAAA TGGGAGAAAT TGGCCAAAT GAAGGGGCTA  
28151 CAGGCCCCAA GTAAGTCCGA AATCTAGTGG AATAGTCAAA TCTTAAAGCT  
28201 CCAAATGAT CTCCTTTGAC TCCACATCAC ACATCCAGCT CATGCTAATG  
28251 CAAGAAGTGG GCTCCCATGG CCTTGGGCAT CTGCACTCCT GTGGCTTTTC  
28301 AGGGTACAGA CCCCCTTCTG GCTCTTTTCA CAGGCTGGCG TTGAGTGTCT

FIGURE 3, page 9 of 42



28351 GTGGCTTTTC CAGGTGCATG GTGCAAGCTG TCGGTGGATC TACTATTCTG  
28401 GGTACTGGAG GATGGTGGCC CTCTTTTCAC AGCTCCACTA GGCAGTGCTC  
28451 CAGTGGGGAC TCTGTGTGAA GGCTCCAACC CCACATTTCC CTTCTGCACT  
28501 GCCCTAGCGG AGGTTCTCCT CAAGGGCTCC ACCCTGCGAG CAAACTTCTG  
28551 TCTGGACATC CAGGCATTTC CATACATCCT CTGAAATCTA GGCAGAGGAT  
28601 CTCAAACCTT AATTCTTATC TTCTGTGTAC CCGCAGACTC AACACCTTGT  
28651 GGAAGCTGCC AGGGCTTGGG GCTTGCACCT TCTGAAGCCA TGGCCTGAGC  
28701 TGTACCTTGG CTCCTTTTAG CCATGGCTGG GATGCAGGGC ACCAAGTCCT  
28751 GAGACTGCAC AAAGCAGCAA GGCCCTGGGC CTGGCCCAGG AAACCATTTT  
28801 TTCCTCCTGG GCCTCTGGGC CTATGATGGG AGGGCCCTTC CTGAAGACCT  
28851 CTGAAGTGCC CTGGAGGCAT TTTCCCATTT GTCTTAGTGA TTAACATTTT  
28901 ACTCCTTGT TCTTATGCAG ATTTCTGCAG CTGGCTTGAA TTTTTCTCTC  
28951 AGAAAACTA TTTTTCTTTT CTGTCACATC ATCAGGGTGC AAATTTGACA  
29001 AACTTTTGTG CTCTGCTTCC TGTGGAATGC TTTGCCACTT AGAAATTTCT  
29051 TCTGCCTGAT ACCCCAAATC ATCTCTCTTA GGTTCAAAGT TCCACAGATC  
29101 TCTAGGGCAG GGGCAAAAAG CCACCACTCT CTTTGCTATA GCATAACAAG  
29151 AGTCATCTTT GCTCCAGTTC CCAACAAGTT CCTCATCTCC ATCTGAGATC  
29201 ATCTCAGCCT GGACTTCATT GCCCATATTA CTGTCAGCAT TTTGGTCAAA  
29251 GCAATTCAAC AAGTCTCTGG GAACCTACAA ACTTTCCAC CTCTTTTGT  
29301 CTTCTGAGCT CTCCAAATTT TTAAGAAGTT CCAAACTTTC CCAGTCTTCT  
29351 TCTGAACCTT CCTAACTGTT CCAACCTCTG CCTGTTACCC AGTTCCAAAG  
29401 TCAGTTCCAT ATTTTTGGGT ATCCTTATAG TAGCACCCAA CTCCTAGTAC  
29451 CAATTTACTG TATTAGTTCA TTCTCACGCT GCTATAAAGA ACCACCTGAG  
29501 AATGGGTATT TTATAAAGGA AAGAGGTTA ATTGACTCAC AGTTTCGCGT  
29551 GGCTGGGGAG GCCTCAGATA ACTTACAGCC ATAGCAGAAA GGAAGCAAA  
29601 CATGTCTCTT CATGTCTCTT AGGAAGAAGA AGTGCTGAGC AAAGAGGGAA  
29651 AAGCCCTATA AAACCATCAT ATCTCGTGAG AACTCACTCA CTATCATGAG  
29701 AACAGCAGCA TGGGGTTGAC CACCCCCCAT AATTCAATTA CCTCCCACCA  
29751 GCTGTCTCCC GTGACACATG GAAATTATGG GAACATAAC TCAAGATGAG  
29801 ATTTGGGTGG GGACACAGCC AAACCATATC ATCTAGGCTG GTATCGAAAT  
29851 CCTGGGCTCA AGCAATCCAC CCACCTTGCC CTACCAAGT GCTGGGATTA  
29901 CAGGCATGAG CCACCATATC TGAAGTGTCT TTTGATTCT TTTGATTTA  
29951 ACCATCCATT GTTCTGCTT CTCTAGATAA CCCTGACTAA TATATAATTG  
30001 GTATGAAGTG ATATCTCATG GCTTTGATTT ATATTTCTTT CATGGCTAGT  
30051 GACTTTTTTT GTACTTTTGG GATATTGTTA TTATTATTAT TATTATTACT  
30101 AGTGTTTATA CTCTTTCAGT AAAAGTGTTA GAAACAATTT TTAAAGGCAG  
30151 AATGTGACCA GAGTTTCCTG TAGTTATATA ACCATCATGG ACCTTCCCTC  
30201 AAGTGCTAAG CCATTAGTGT TACTCATGTC ACTCCAAATG TCAGCTTGTT  
30251 TTCTTCCATT TCACCTGTCT TTTGTGTCCC AAACCTGAAT TCATGGGAAA  
30301 AACATCTGAA TGGTGCTTAA TATGGTTTGG ATATTTGTCC CCTCCAAATC  
30351 TCATGTTGAA ATATGACCTC CAGTGTTGGA AGTAGGGACT ACTTGGGTCA  
30401 CGAGAGTGGA TCCTTCATTA ATGGCTTGGT AATAAGTGAA CTCTATTAGT  
30451 TCATGAAAGC TGGTTGTGTA TAAGAGCCTG GCATCTCAT TCTCTTGTC  
30501 TTCTCTCACC ATCTGACACA CTTGCTCACC TTTTTCTTC AGCCATGAGT  
30551 AAAAGCTTCC TGAGGTCTCA CCAGAACTG AGCAGATGTT GGTGCCATGC  
30601 TTGTACAGTC TGTAGAACTG TGAGCCAAAT AAGCCTCTTT TCTTTATAAA  
30651 TTACCGAGTC TCAGGTGTTT GTTTAAAACA ACACAAAACA GACTAACACA  
30701 GTGTTGATTG AAACAGCTGT GACTGGGTCA TCAGGGTGTA AGAGAGGAGT  
30751 CACTGAGTTG AAATATAGCC TCCTACTTAC ACCTGTTCAG TAGAAGCTGT  
30801 AGATATGAAG TAGCTGAAGC AGGCATTCCC TCTGAAACAT GTGTTTCACA  
30851 TATGTCATAA TTATCTTCTG CTCTCATTTT TCTTTTAGGC TTTTGTCTCC  
30901 ATCTCATTTT CCCTGTTTAC TCTCATTTT ATATCTTTAC ATTTCTTTCT  
30951 CCAGAATTGT TCAGAAGCTT GGAACCTTC ACTCCAGTTA TTCTTTGACT  
31001 ATGCAATTTG TTTCTGTGCT TCATGGCACT TATGGTTTGT AATCCTTGAC  
31051 TTGTTTGTAT AGCTCAGTGG TTAGGAGTAC AGTTTGGAGT TAGAATGCCT  
31101 GGGTTGAAAC TCTTAATTCT ACTCTACTTA CTAGTCTTGT GACTATAACA  
31151 AAATCTTAG CTTCTCTTTG TCTGTAAAAT GGAGAGTATA GTAAATACAT  
31201 GGGCTTGTTT CAGGATTA ATGAGTTAAT ATGTGAAATA CTAGAACAA  
31251 TGCCTGGCAA ATGCTCAATG AATATTGAGT ATTGCTTGCT TTTGTTTGT  
31301 GCCATGCCCT TGTTCCTCAC TGAGGGCACA GACCATGTGT ATCTGGTTAA  
31351 CAGTTCTATG TCCACCACGT TGCAATAATG GACTCTCAGA AAATATTGAA  
31401 GAATATGTTA AAGAATGAGT AGAATTATGC TACTGAAAAG GGTGAGTGGA  
31451 AGGTAGGTAG GGGAAAGGAC ATATACAGCC CTGGAGGCAG CATATATGGG

FIGURE 3, page 10 of 42



31501 GAATGGGTCA CACAGTGTTC CTTGGTACTC TCTAGACCAT AGTGGGCCAC  
31551 CTCTTAGCTA GTGGCTATG GATTATTTCA GCAGTCTGTT GGAAACATCC  
31601 ATGAATATGA TAATAATGAC CCATTGTGTG GTTCTAAGAA AAAGGACAAC  
31651 TACAATACTA GACAATAATA GTATGTAAGT TAGGAGGGAA GGGGATGATT  
31701 TGTATTAAAC TGTTCTAAAA TTCTTACCTT ATTTAGGATG ATGGGGTCAG  
31751 ACATTAACCT TAGACTTTGT TATATATATG TGGTAAATTT TCAAGGTAAA  
31801 CCATTGAAAC TGTAGTAGTT GAGTATATAA CTTCCAAATC AGGGGGGAAA  
31851 GAAATGGAAT AAGAAAATAA ATACATAAAC ATAAGATTGA AACAATCCAA  
31901 TGAAGAGTAG AGAGAAGAGG GAAAAACATA GAAAGAATGA GATAATTAGA  
31951 AAGCAATAGG TAAGATGTGA GAAATAAATT CAAGTACAGT AAAACTCCAC  
32001 TAAAATGTGC CCTGCAGTAA TGTTGGGGCA TGATTTCCTT TCATCCCCAT  
32051 TCTCAAATGG GGCAGCGTAA ATAGCGTTCT TATCCTGTTT CCCTGGGGGT  
32101 TTGAGGTGGG TGACGAGTAA GTTAGAAGAT AATCACCTTC TGATCAGTTA  
32151 GGACTTTCTC AGTTTAGTCT TCAATTAATA AAAATTAATG TAAATTTTCAT  
32201 CAGAAGGCAG AGATTGTCAG ATGAAAGAAC AAGCAAAATA AAAGTCTTAC  
32251 TGAAAAAAG CTGGGGTAGC TATGTTAATA TCAACTGTGA ATTATTATTA  
32301 ATAATCTATT AATAATAGAT TATATAGTAA AAACATTAAT AAAAATAGAG  
32351 TGCTACTACA TTTTAAAAAT CAGTATGAGG ATATACAATT TTTAAGCTGG  
32401 TTGATAAAAT AATTGGGATT AATTGGGAAA TCCATCATAG TGGTGAGAGA  
32451 TTTTAAACACA ATTCTTCCTG TATTTGATAG GTCAAGCAGA GAAAACTTTT  
32501 AGTGAAGACA AAAACTTCTA AATACATAAG CTTGATTAA TGGGCATGTA  
32551 ATAGGACCTA GCATCAAAAA ATTAGAAAAA ATATTTTTTC TTAGGTATTT  
32601 ATGGAACATG TATAAAAAAT GATTTCTGTAG TAGGCCATAA AGCCAGGTTC  
32651 AACACATTTT AAAGAAGTGG TATCACAAGA ACTGCTTTCT CTGACCACTA  
32701 TGCATTAAAA TAGAAGTTAA TTACAGACAT AAATTATAAA AATGCCAATA  
32751 TTTTAAAGTG TGATATACAC TTCTCACTT ATGGGTCAAA GGAAATCGTA  
32801 AGTGGAATTT CAAGGACACG TTGACTTGAA AACATTAAAA CTTATGGAAT  
32851 ATTTCTAAGA TGGAAGTTGT ATGAATTTTA TAGTCTGAAA GCTTTTATTA  
32901 GAAAAGAATT AAGTCTGAAA ATTAATGTGC TAAGTTAGGG GAGAGAAAAT  
32951 GGAATAATCT CGAAGAAGGT AGGAGGAAGG AGATAATAAA GAATATATAG  
33001 CAAAGATGCA GTAACAGGAT CAACAAAGCC AGAACTGTT GGAAAAGACA  
33051 AGCCTCTGGA AAGATTGATG AAGAAAAAAG AGAAATGAGA TGTAAATAAA  
33101 TCATGTTTCTG TATATAATAG GCACATAAGG ACTTTTAAAA AACTAATAAA  
33151 ATAATATGAA TCATTAATGC CAATAAATTT GAAAACAGAC AAAGTAGGTG  
33201 AATTTCTAGA AAAATATAAC TTAGTGGGAC TGAATGAAGA AGCAACAGCT  
33251 TATAGTACCT AAGCAATTGA AGAGATTGGG TCAGTAATTT AAAATTTTCT  
33301 CATAAACAAA ACCTTAGCCC CAGATGGTTC TTGCAAATGA TTAAAGAACA  
33351 GATGTACAAA CATTTCCAGA GTGTAGAAGT ACCTGTCTCT ATCCTTTCTA  
33401 GGAGATCATT ATAACACCAA AAGCAGACAG TATATGAAAC AGGGAAATTA  
33451 GAGGCCAAGA TACCTATGAC TTATATGTAA AAATTTAAAG AAAATATTAG  
33501 CAACTGAAT CAGCCATTTT AAAAAATATA CCACAATCAA TGCATTCTA  
33551 AGAGCAGCTT AACAAAAATTT GTTAGAAGGC ATTAAGAAG ACTCAGTATA  
33601 GAAAAGATGT ACCTTCTCTC CAAATTGGTG ATAGAGATTC AATGCCATTA  
33651 AAAAAACCCA CCTGGTTTTT TTGAGGAACT TGTCAGCTG AGTCTCAAAT  
33701 TTATATCAAA GAGCAAGGC CTAAGAATAT CCAGGACATT CCTGAAGAAC  
33751 TGTAAGGAGC CAGGGGCCTG CCCTATCAGA TACCAAGGGT TGTTATTAAAG  
33801 CCATAACCAA GTCAGTGCTG TTTCTACAGA AACAGACAAG TTAACAAGTG  
33851 AAACATAATA GAGAGCCAG AAACAGACCC ATCCATATTT TGGATTTGTC  
33901 ACGTGAAAGA AGTAGCTTTG CAAAACCTTG GGGAAAGGAG AGTGTGTGCA  
33951 ATAGATGATG CTCGTGCTCA TGCAGACAAA AAGGAAATTG GGATACCTGC  
34001 CTCTTACCGT ACACAAACAC CAACCTAAAC GTGAAAGTTA AACTATAACA  
34051 GCTTGAGGTG GTGGGGAGA AATATCTTTA TCTCAGTGTA GGGAGAATT  
34101 TATTTTAAAA AGAAGACACA AAAGGCCATA CATAGGAATG AAAAGATTGA  
34151 ATTCAGCTGC ATTAATAAGA TTAAATTCAG CTGCGTTAAA ATCAAGAGCA  
34201 TCTGTACTTG GACAGCATAG AGTGGAAAGA CAAAGAGAAG GTATTTGCCA  
34251 GCTTATAACT TGAAGGATTA GAATGAATGA TATAAAGAAC TATGTAAATA  
34301 AGAAAAAGAC ATACAACCGG TTAGAAAAAC GGGCAAAGAC ATGAACAGCA  
34351 TATTTACAGT GAAGGAAACA GCGGTAGCAA ATGAACATGG TAAGAGATGC  
34401 TCAACACGTT TAGTAATTTG AAGGGAAATG CAAGTTATAC CCACAGCAAG  
34451 ACTATCTTAT CTAGGAAGTT TGTCAATACC CTAAATGTTT TGTGGTTTTA  
34501 AGCTACAGAG TTTGTAATTC ATTTATTTAT TCAATAAATA CTCAGTGGCA  
34551 GGCACTGTTT TAGAAACCTT GGTATATACT TTGAATGAAA TTAACAAAAA  
34601 TCCTTGCCCT GTGGAGGATG CTTATGTGTG GGGAGTTGGG TGGTGGGGTC

FIGURE 3, page 11 of 42

34651 AAACAACAAT TACATTAAAA TAGAAAATAG TGACATAAAT AAACCTATAA  
34701 ATATTGCAAC CCAGAGTTAT ATTATAAATG TAAGTAGTGA CTAGGACTCT  
34751 CATGCAGATA TACCTCTGTG CTGGGACAAA TGAAAGTTTA AGTGTAATTT  
34801 CCCATATGCA AGTCAAAATA AAAAGTGACA CTAGAAAACA CAATAATGAA  
34851 TATCTGAAAA TTGCATTTTA TTTGACTGCC ATCCTTTTGC ATCATTTTCA  
34901 TACTAATTAT AGAATAAAAT TTGTAGGATG CACCAAAGCT TTTTITAGAG  
34951 ACATCCATTA ATTCATAAAA TAAATGAGCA CCTTCTTTGT GCCAGCAGCT  
35001 GTAAGAGGTG GCCCAAGGAA GGGAAATAAA CAGTCAAAAT CCTGGTACAC  
35051 TCAGAGTTTC TCTTAGGAGA AAACAGATAC AAATGGCATT AATTACCAAG  
35101 AAACCTGTAA AACAAGCCAA ATATTAATGA TAAATATTTG AGTACAGTAT  
35151 GTTAATTTTA AGATTGAAAA TGAGGTGCCA GGATTCTTGA AGACTCAAAG  
35201 GCGAAGATGG CATTGAATGA ACAGCTCTGG TCTACAGCTC CCAGCGTGAG  
35251 CGACGCAGAA GACGCATGAT TGCTGCATTT CCATCTGAGG TACCGGGTTC  
35301 ATCTCACTAG GGAGTGCCAG ACAGTGGGCG CAGGTCAGTG GGTGTGTGCA  
35351 CCGTGCGCGA GCTGAAGCAG GGCGAGGCAT TGCCCTCACTC GGGAAAGTGCA  
35401 AGGGGTGAGG GAGTTCCTTT TCCTAGTCAA AGAAAGGGGT GACAGATGGC  
35451 ACCTGGAAAA TCGGGTCACT CCCACCTGAA TACTGCACTT TTCTGACGGG  
35501 CTTAAAAAAT GCGGCACCAG GAGATTATAT CCTGCACCTG GCTCGGAGGG  
35551 TCCTACACCC ACGGAGTCTC GCTGATTGCT AGCACAGCAG TCTGAGATCA  
35601 AACTGCAAGG CGGCGGCGAG GCTGGGGGAG GGGCACCCGC CATTGCCAG  
35651 GCTTGCTTAG GTAAACAAAG CAGCCGGGAA GCTCAAACTG GGTGGAGCCC  
35701 ACCACAGCTC AAGGAGGCCT GCCTGCCTCT GTAGGCTCCA CCTCTGGGGG  
35751 CAGGGCACAG ACAACAAAA AGACAGCAGT AACCTCTGCA GACTTAAATG  
35801 TCCTGTCTG ACAGCTTTGA AGAGAGCAGT GGTTCCTCCA GCACGCAGCT  
35851 GGAGATCTGA GAACGGGCAG ACTGCCCTCT CAAGTGGGTC CCTGACCCCT  
35901 GACGCCCCGAG CAGCCTAACT GGGAGGCACC CCCCAGCAGG GGCACACTGA  
35951 CACCTCACAC AGCCGGTTAC TCCAACAGAC CTGCAGCTGA GGGTCTGTGTC  
36001 TGTTAGAAGG AAAACTAACA AACAGAAAGG ACATCCACAC CAAAAACCCA  
36051 TCTGTACATC ACCATCATCA AAGACCAAAA GTAGATAAAA CCACAAAGAT  
36101 GGGGAAAAAA CAGAGCAGAA AAACCTGAAA CTCTAAAAAG CAGAGTGCCT  
36151 CTCCTCTCTC AAAGGAACGC TGTTCTCTAC CAGCAACGGA ACAAGCTGG  
36201 ATGGAGAATG ACTCTGACGA GCTGAGAGAA GGCTTCAGAC GATCAAATTA  
36251 CTCTGAGCTA TGGGAGGACA TTCAAACCAA AGGCAAAGAA GTTGAAAACT  
36301 TTGAAAAAAA TGTAAGAAGT TGTATAACTA GAATAACCAA TACAGAGAAG  
36351 TGCTTAAAGG AGCTGATGGA GCTGAAAACC AAGGCTCGAG AACTACATGA  
36401 AGAATGCAGA AGCCTCAGGA GCTGATGCGA TCAACTGGAA GAAAGGGTAT  
36451 CAGCGATGGA AGATGAAATG AATGAAATGA AGCGAGAAGG GAAGTTTGA  
36501 GAAAAAAGAA TAAAAAGAAA CGAGCAAAGC CTCCAAGAAA TATGGGACTA  
36551 TGTGAAAAGA CCAAATCTAT GTCTGATTGG TGTACCTGAA AGTGACGGGG  
36601 AGAATGGAAC CAAGTTGGAA AACACTCTGC AGGATATTAT CCAGGAGAAC  
36651 TTCCCCAATC TAGCAAGGCA GGCCAACATT CAGATTGAGG AAATACAGAG  
36701 AACGCCACAA AGATACTCCT TGAGAAGAGC AACTCCAAGA CACATAATTG  
36751 TCAGATTAC CAAAGTTGAA ATGAAGGAAA AAATGTTAAG GGCAGCCAGA  
36801 GAGAAAGGTC GGGTTACCCT CAAATGGAAG CCCATCAGAC TAACAGCGGA  
36851 TCTCTTGGCA GAAACTCTAC AAACCAGAAG AGAGTGGGGG CCAATATTCA  
36901 ACATTCTTAA AGAAAAGAAT TTTCAACCCA GAATTTTATA TCCAGCCAAA  
36951 CTAAGCTTCA TAAGTGAAGG AGAAATAAAA TCCTTTACAG ACAAGCAAAT  
37001 GCTGAGAGAT TTTGTACCCA CCAGGCCTGC CCTAAAAGAG TTCCTGAAGG  
37051 AAGTGCTTAA CTTGGAAAGG AACAATCAGT ACCAGCCGCT GCAAAATCAT  
37101 GCCAAAATGT AAAGACCGTC GAGACTAGGA AGAAACTGCA TTAACAAACG  
37151 AGCAAAATAA CCAGCTAACA TCATAATGAC AGGATCAAAT TCACACATAA  
37201 CAATATTAAC TTTAAATGTA AATGGACTAA ATGCTCCAAT TGAAAGACAC  
37251 AGACTGGCAA ATTGGATACA GAGTCAAGAC CCATCAGTGT GCTGTATTAA  
37301 GGAAACCCAT CTCACATGTA GAGACACACA TAGGCTCAAA ATAAAAGGAT  
37351 GGAGGAAGAT CTACCAAGCA AATGGAAAAC AAAAAAGAC AGGGGTTGCA  
37401 ATCCTAGTCT CTGATAAAAC AGACTTTAAA CCAACAAAGA TCAGAAGAGA  
37451 CAAAGAAGGC CATTACATAA TGGTAAAGGG ATCAATTCAA CAAGAAGAGC  
37501 TAACTATCCT AATATATAT GCACCAATA CAGGAGCACC CAGATTGATA  
37551 AAGCAAGTCC TGAGTGACCT ACAAGAGAGC TTAACCTCCC ACACATTAAAT  
37601 AATGGGAGAC TTTACACCCC CACTGTCAAC ATTAGACAGA CCAATGAGAC  
37651 AGAAAGTCAA CAAGGATACC CAGGAATTGA ACTCAGCTCT GCACCAAGCA  
37701 GACCTAATAC ACATCTACAG AACTCTGCAC CCCAAATCAA CAGAATATAC  
37751 ATTTTTTTCA GCACCACACC ACGGCTATTG CAAATTTGAC CACATACTTG

FIGURE 3, page 12 of 42

37801 GAAGTAAAGC ACTCCTCACC AAATGTAAAA GAACAGAAAT TATAGCAAAC  
37851 TATCTCTCAG ACCACAGTGC AATCAAACCTA GAACTCAGGA TTAAGAATCT  
37901 CACTCAAAAC CGCTCAACTA CATGGAAACT GAACAACCTG CTCCTGAATG  
37951 ACTACTGGGT ACATAACGAA ATGAAGGCAG AAATAAAGAC GCTCTTTGAA  
38001 ACCAACAAAGA ACAAAGACAC AACATACCAG AATCTCTGGG ACGCATTCAA  
38051 AGCAGTGTGT AGAGGGAAAT TTATAGCACT AAATGCCAC AAGAGAAAGC  
38101 AGGAAAGATC CAAAATTGAC ACCCTAACAT CACAATTAAA AGAACTAGAA  
38151 AAGCAAGAGC AAACACATTC AAAAGCTAGC AGAAGGCAAG AAATAACTAA  
38201 AATCAGAGCA GAACTGAAGG AAATAGAGAC ACAAAAAACC CTTCAAAAAA  
38251 TTAATGAATC CAGGAGCTGG TTGTTTTTGA AAGGATCAAC AAAATTGATA  
38301 GACCGCTAGC AAGACTAATA AAGAAAAAAA GAGAGAAGAA TCAAATAGAC  
38351 ACAATAAAAA ATGATAAAGG GGATATCACC ACCAATCCCA CAGAAATACA  
38401 AACTACCATC AGAGAATACT ACAAACACCT CTATGCAAAT AAAC TAGAAA  
38451 ATCTAGAAGA AATGGATAAA TTCCTCGACA CATAACCCCT CCCAAGACTA  
38501 AACCAGGAAG AAGTTGAATT TCTGAATAGA CCAATAACAG GATCTGAAAT  
38551 TGTGGCAATA ATCAATAGCT TACCAACCAA AAAGAGTCCA GGACCAGATG  
38601 GATTACAGC CGAATTCTAC CAGAGGTACA AGGAGGAAT GGTACCATT  
38651 CTTCTGAAAC TATTCCTAATC AATAGAAAAA GAGGGAATCC TCCCTAACTC  
38701 ATTTTATGAG GCCAGCATCA TCCTGATACC AAAGCCAGGC AGAGACACAA  
38751 CAAAAAAGA GAATTTTAGA CCAATATCCT TGATGAACAT TGATGCAAAA  
38801 ATCTCAATA AAATACTGGC AAAGTGAATC CAGCAGCACA TCAAAAAGCT  
38851 TATCCACCAT GATCAAGTGG GCTTCATCCC TGGGATGCAA GGCTGGTTCA  
38901 ATATACGCAA ATCAGTAAAT GTAATCCAGC ATATAAACAG AACCAAAGAC  
38951 AAAAACCACA TGATTATCTC AATAGATGCA GAAAAAGCCT TTGACAAAAT  
39001 TCAACAACAC TTCATGCTAA AAAGTTTCAA TAAATTAGGT ATTGATGGGA  
39051 TGTATCTCAA AATAATAACA GCTATCTATG ACAAACCCAC AGCCAATATC  
39101 ATACTGACTG GGTAAAAACT GGAAGCATTC CCTTTGAAAA CTGGCACAAG  
39151 ACAGGGATGC CCTCTCTCAC CACTCCTATT CGACATAGTG TTGGAAAGTTC  
39201 TGGCCAGGGC AGTTAGGCAG GAGAAGGAAA TAAAGGGTAT TCAATTAGGA  
39251 AAAGAGGAAG TCAAATTGTC CCTGTTTGCA GACGACATGA TTGTATATCT  
39301 AGAAAACCCC ATTTGTCTCAG CCCAAAATCT CCTTAAGCTG ATAAGCAACT  
39351 TCAGCAAAGT CTCAGGATAC AAAATCAATG TACAAAAATC ACAAGCATTC  
39401 TTATACACCA GCAACAGACA GAGAGCCAAA TCATGAGTGA ACTCCCGTTC  
39451 ACAATTGCTA CAAAGAGAAT AAAATACCTA GGAATCCAAC TTACAAGGGA  
39501 TGTGAAGGAC CTCTTCAAGG AGAACTGCAA ACCACTGCTT AATGAAATAA  
39551 AAGAGGATAC AAACAAATGG AAGAACATTC CATGCTCATG GGTAGGAAGA  
39601 ATCAGTATCG TGAAATGGC CATACTGCCC AAGGCAATTT ACAGATTCAA  
39651 TGCCATCCCC ATCAAGCTAC CAATGACTTT CTTACAGAAA TTGGAAAAAA  
39701 CTACTTTAAA GTTCATATGG AACCAAAAAA GAGCCCGCAT TGCCAAGTCA  
39751 ATCCTAAGCC AAAAGAACAA AGCTGGAGGC ATCATGCTAC CTGACTTCAA  
39801 ACTATACTAC AAGGCTACAG TAACCAAACC AGCATGGTAC TGGTACCAA  
39851 ACAGAGATAT AGACCAATGG AACAGAACAG AGCCCTCAGA AATAACGCCG  
39901 CACATCTACA ACTATCTGAT CTTTGACAAA CCTGAGAAAA ACAAGCAATG  
39951 GGGAAAGGAT TCCCTATTTA ATAAATGGTG CTGGGAAAAA TGGCTAGCCA  
40001 TATGTAGAAA GCTGAACTG GATCCCTTCC TTACACCTTA TACAAAAATC  
40051 AATTCAAGAT GGATTAAAGA CTTAAACGTT AGACCTAAAA CCATAAAACC  
40101 CCTAGAAGAA AACCTAGGCA TTACCATTCA GGACATAGGC ATGGGCAAGG  
40151 ACTTCATGTC TAAAACACCA AAAGCAATGG CAACAAAAGC CAAAATTGAC  
40201 AAATGGGATC TAATTAAACT AAAGAGCTTC TGCACAGCAA AAGAAACTAC  
40251 TATCAGAGTG AACAGGCAAC CTCCAAAATG GGAGAAAATT TTTGCAACCT  
40301 ACTCATCTGA CAAAGGGCTA ATATCCAGAA TCTACAATGA ACTCAAACAA  
40351 ATTTACAAGA AAAAAAACAA ACAACCTAT CAAAAAGTGG GTGAAGGACA  
40401 TGAACAGACA CTTCTCGAAA GAAGACATTT ATGCAGCCAA AAAACACATG  
40451 AAAAAATGCT CACCATCACT GGCCATCAGA GAAATGCAAA TCAAAACCAC  
40501 AATGAGATAC CATCTCACAC CAGTTAGAAT GGCAATCATT AAAAAGTCAG  
40551 GAAACAACAG GTGCTGGAGA GGATGTGGAG AAATAGGAAC ACTTTTACAC  
40601 TGTGGTGGG ACTGTAACT AGTTCAACCC TTGTGGAAGT CAGTGTGGCA  
40651 ATTCTTCAGG GATCTAGAAC TAGAAATATC ATTTGACCCA GCCATCCCAT  
40701 TACTGGGTAT ATACCCAAAG GACTATAAAT CATGCTGCTA TAAAGACACA  
40751 TGCACATGTA TGTTTATTGT GGCATATTTC ACAATAGCAA AGACTTGGA  
40801 CCAAGCCAAA TGTCCAACAA TGATAGACTG GATTAAGAAA ATGTGGCACA  
40851 TTTACACCAT GGAATACTAT GCAGCCATAA AAGATGAGTT CATGTCTTTT  
40901 GTAGGGACAT GGATGAAATT GGAATCATC ATTCTCAGTA AACTATCACA

FIGURE 3, page 13 of 42

40951 AGAACAAAA ACCAAACACC GCATATTCTC ACTCATAGGT GGGAATTGAA  
41001 CAGTGAGAAC ACATGGACAC AGGAAGGGGA ACATCACACT CTGGGGACTG  
41051 TTGTGGGGTG GGGGGAGGGG GAGGGATGGC ATTGGGAGAT ATACCTAATG  
41101 CTAGATGACG AGTTAGTGGG TGCAGCGCAC CAGCAAGGCA CATGTATACA  
41151 TATGTAACATA ACCTGCACAT TGTGCACATG TACCCTAAAA CTTAAAGTAT  
41201 AATAATAAAA AAAAAAGACT CAAAGGCACA GTCAGTACA GTTTGATTTT  
41251 TTATAATAGC TGTTAATTTT CCTAACTTCG AGGAAGTTGA TAGCATGTTT  
41301 TGAGTATATT TCAAACTAC ATTCAAATGT TGCAATAGAA CATTAAAGAAT  
41351 TATCTTCATG ATCCACTAAG TGCATGAAAA AAATGGATAA TGAATCTATT  
41401 CATTACCATC GTTTAATATT TTATCTTCAA GTTTTTGTGT TTTGTAGCTC  
41451 ATTGGCAGAG TTTGACAGAG TGCTGAAAGT ATTCTTTAGT GAGCTGGCTG  
41501 TAATTTTTGG GCCCATTTTT ATCTAGATAA TTAAACTAT CTGACAGGAC  
41551 CATAAAATGC TTGCTGCCAT TTCCAACAAC CTATATTTGT GGATGGGGTT  
41601 TTTTAATTTA ATGAGAATAT TATGTTAGAA AAGAACTGT CATTCTGTAA  
41651 AGTGGCCAAT AATGTTAGTT TTATTTATCA ATTTAGTTTT GTACTTTGAT  
41701 CATTTTTTTA AAATTTTCAGC ATTGATGTTG ATGGGACAAAT GACAGTGGAC  
41751 TGGAATGAAT GGAGAGACTA CTTCTTATTT AATCCTGTTA CAGACATTGA  
41801 GGAAATTATC CGTTTCTGGA AACATTCTAC AGTAAGTCTA CTTTATGTAT  
41851 TTATACTTAT TTGGAGCTAT AAACCATAGG TACAGTTATC ACCCAAGAAC  
41901 ACTCTGTAAC ACTTATGGGC CAGGATACCT GAGTCCCAGT AGCTCCTTAA  
41951 CCTGTAGAGT TCTATTTATT CTATTAGGCA TAGATTTATA GAGTATTAAA  
42001 CAAAAAATAA CAGCTCTCCC TCTCCCTCTC CCTCTCTCTC CCCCTCCCA  
42051 CGGTCTCCCT CTCCCTCTCT TCCACCGTC TCCCTCTGAT GCCGAGCCAA  
42101 AGCTGGACTG TACTGCTGCC ATCTCGGCTC ACTGCAACCT CCCTGCCTGA  
42151 TTCTCCTGCC TCAGCCTGCC GAGTGCCTGC GATTGCAGGC GCGCACC GCC  
42201 ACGCCTGACT GTTTTCTGTA TTTTTTGGT GGAGACGGG TTTCTGCTATG  
42251 TTGGCCGGGC TGCTCTCCAG CTCCTGACCG CGAGTGATCC ACCAGCCTCG  
42301 GCCTCCCGAG GTGCTGGGAT TGCAGACGGA GTCTCGTTCA CTCAGTGCTC  
42351 AATGGTGCCC AGGCTGGGGT GCAGTGGCAT GATCTCGGCT CGCTACAACC  
42401 TCCACCTCCC AGCCGCTGC CTTGGCCTCC CAAAGTGCCA AGATTGCAGC  
42451 CTCTGCCACG CCGCCACCCC GTCTGGGAAG TGAGGAGCGT CTCTGCCTGG  
42501 CCGCCCATCG TCTGGGATAT GAGGAGCCCC TCTGCCTGGC TGCCAGTCT  
42551 GGAAAGTGAG GAGTGTCTCT GCCCGGCCGC CATCCTGTCT AGGAAGTGAG  
42601 CGTCTCTGCC CGGCCGCCCA TCGTCTGGGA TGTGAGGAGC CCCTCTGCCT  
42651 GGCTGCCCAG TCTGGAAGT GAGGAGCGCC TCTTCCCGGC CGCCATCCCA  
42701 TCTAGGAAGT GAGGAGCGTC TCTGCCCGGC CGCCCATCGT CTGAGATGTG  
42751 GGGAGCGCCT CTGCCCGCC GCCCGTCTG GGATGTGAGG AGCGCCTCTG  
42801 CTCGGCCGCC CCGTCTGAGA AGTGAGGAGA CCCTCCGCC GGCAGCGGCC  
42851 CCGTCTGGGA AGTGAGGAGC GTCTCCGCC GGCAGCCACC CTGTCCGGGA  
42901 GGGAGGTGGA GGGGTGAGCC CCCCGCCCG CCAGCCACCC CATCCGGGAG  
42951 GTGAGGGGTG CCTCTGCCCG GCCGCCCTA CAGGGAAGTG AGGAGCCCCT  
43001 CTGCCCGGCC ACCACCCCAT CTGGGAGGTG TACCCAACAG CTCATTGAGA  
43051 ACGGGCCATG ATGACAAATG CGGTTTTGTG GAATAGAAAA AGGGGAGAGG  
43101 TGGGGAAAAG ATTGAGAAAT CGGATGGTTG CTGTGTCTGT GTAGAAAGAG  
43151 GTAGACATGG GAGACTTTTC ATTTTGTCT GTACTAAGAA AAATTCTTCT  
43201 GCCTTGGGAT CCTGTTGATC TATGACCTTA CCCCCAACCC TGTGCTCTCT  
43251 GAAACATGTG CTGTGTCCAC TCAGGGTTAA ATGGATTAAG GCGGGTGCAA  
43301 GATGTGCTTT GCTAAACAGA TGCTTGAAGG CAGCAGGCTC GTTAAGAGTC  
43351 ATCACCACCT CCTAATCTCA AGTACCCAGG GACACAAACA CTGCGGAAGG  
43401 CCGCAGGGTC CTCTGCCTAG GAAACCAGA GACCTTTGTT CACTTGTTTA  
43451 TCTGCTGACC TTCCCTCCAC TATTTGCTG TGACCTTGCC AAATCCCCCT  
43501 CTGCGAGAAA CACCCAAGAA TGATCAATTA AAAAAAATA AAAAAAACA  
43551 ACCCAAGACT GCATAAATGT CCATTCTGAA AACTTGAAG AAGTACCACC  
43601 TTGATGAATA AGCTGTCTAG CTTTTATTGG CATTAAAGTA TTCTGCCATA  
43651 GGAAGTGTA AAAGTTGTAG GCTTTTACTT TTTATAGGTA CTATATTGTC  
43701 CAAATAATCT CAGCACCTCA TGSTTGCTAA GGATCTGTGT CCTTGTGTTG  
43751 TCAGATTATG TTTATCTCTG GCATAAGGCA CTTAACAATA TTCATTAAAG  
43801 GTTACAGAAAT CTTTTGCTT CATCTGCTTA GCATTTTATA CCAGTTTGT  
43851 TTCCACCAAA CTTTCAAATT TTGATTGTTT CATTAATATT CTGCATACTG  
43901 ATGTAAACCA AGTTCTATTA TTGTGCAATC TGCTCCTGAA ACCCTTAGGA  
43951 ACTCTCTGAA GGAGTTTTAT TTATTTTTTG TTTTGTGTTT TGTTTTGTT  
44001 TTGTTTTTTT GAGACGGAGT CTTGCTCTGT TGCCAGGCT AGAGTGCAGT  
44051 GGTGCGATCT CGGCTCTCTG CAACTCGGC CTCCGGGGT CACGCCATTC

FIGURE 3, page 14 of 42

44101 TCCTGCCTCA GCCACCGGAG TAGCTGGGAC TACAGGCACC CACCACTGCG  
44151 CCTGGCTAAT TTTTTTTGTA TTTTGTAGTAG AGACGGGGTT TCACCGTGT  
44201 AGCCAGGATG GTCTCGATCT CCTGACCTTG TAATCCGCCC GCCTCGCCTC  
44251 CCAAAGTGCT GGGATTACAG GCGTGAGCCA CTGTGCCCGG CCTTTTTTTT  
44301 TTTTTTTGCT TTATGGGCTT GTCTTCTACA CTTTCAGATTT GACTAAATTA  
44351 AATATGCATT AAATGAAGTC AGGAGTTCAC ATTGCCACTA GTAACAATGC  
44401 CTAAGCTTAC ATAAAGCATT ATAAAATTGT TGGTGATTAG TGCCTTCTCA  
44451 GCTATGAGTA TAAGATAATA TTATACTAGT AGTTCAGTTG CCTAGATAAA  
44501 TTGTACACTA TGTGAAGTTT TATTTACATA ATTCTTACGG TATTTTTTAA  
44551 GGTAGTTGAT AACAGTTGAG ACTACAATTG TATCTCCATT TTATTGATAG  
44601 TAAATGAAG GAAGGGGAGG TTAATAACAT AGGAGAGCTC CTCCCGTTG  
44651 CACTCTTGCC TTGAAAAATT TTTCTGCCAA AACAATTTAG ATAATAGAAT  
44701 TGTAAAAATA TTATTATAGA ATTGTTTCTC TCAAACATA GTAATGTAGA  
44751 ATAGGTTGAA GGGGTGATGA TTTGAAACAA TACCTCTCCA TTAGCTAAAT  
44801 TTTATATAGA ATCTATTGCA TGTTTTAAAT GATAAGTCAG ATTTATAAAA  
44851 ATATTTTTAT AACAGTAGG AAATGAGTTT AGGGGTATTC ACATACAGTT  
44901 TTAATTTTTA TTTACATATT TAAAACATAT CATGGTATAA ATATGATGTG  
44951 GATATAAATT TGAGATAAAG GAAGTATTGT TTAAGAATTG ATGAACTAAT  
45001 TTCTTAAAAA ATGTCATCAC CAGTTGGTTT TCTAGCCTTA TGAAAAATGG  
45051 TTGCAATAAA AAAGATTGAC TATGATAAAA TGCTGCCCTT TCATTTTAAC  
45101 CTAGACCAAG AGAAAAACATA CTGTGAATCT ATGATGAATG AAAGAAAGTT  
45151 GTAACGTGTG GTTTTGTATA TTTGTAATTA CTGTTTATTT TCATTTCTTG  
45201 TGAACGTGATA CTGTACTTTG TTCATTGTGA GTAGACAACT TATAATCTAT  
45251 GTACTCAAAAT TGGTTTAGTA TAAATTCTAG GGAATGAAGT TCATATTAAC  
45301 TGTAAAAATA CATGATTGTT CTCTAAACA AAACGTCTTC TGGGATTATT  
45351 TTTAACTAAG GCGCATGGG ATCTTTTTTT CATTTTTACA GGGAAATTGAC  
45401 ATAGGGGATA GCTTAACTAT TCCAGATGAA TTCACGGAAG ACGAAAAAAA  
45451 ATCCGGACAA TGGTGGAGGC AGCTTTTGGC AGGAGGCATT GCTGGTGCTG  
45501 TCTCTGAAC AAGCACTGCC CCTTTGGACC GTCTGAAAAT CATGATGCAG  
45551 GTGAGCTTTA TTATCGTGTG TCCAGGTTTG CCCTAAATAT TCTAAAAACA  
45601 TGAGAAATGT GGTGCTTTGA AAAAGAAGTT TTAATTTTC TCAGTAATAA  
45651 TCTTTTATAC CCTAAAAAAT AAATCTATTT TGTTGCTGTT AACTCTAAAT  
45701 TCAAGTCCATG TAAGTATGGC AGTGTACCAA ACCTTAAATT GTTAGTACAT  
45751 GTGTGTAATG AACTTTTAAT CTTTGGCATT CTATGACTAT TCAAACATTT  
45801 AATTCAAAAA ATATCTCTAG CTATTGTTGT AGGATTCTCC TGATTTATAG  
45851 TTTCTTCTT TTTAATATAC TTTATCAAAA GTAAAGTATT TTTGAAATCT  
45901 AGACTCTTAG AGCAGCAATG TAATTTTGAA AATTATTCTA AAGCTGAGGT  
45951 TAGCAGAAAA AGATCTGGCT TTATAGACTG ACTTTGCTAT TTAGTAGCAG  
46001 TGTAGCATTG GGCTGGCCAG AGTGGAAAGA GGAATGGAA AAGAATTAAT  
46051 ATGTATTGTC TCACTGTGGT AACCCAGTTA ATCCTTGCG CAGCCAGTG  
46101 AAGTAGGTAT TTTATCATTT TTCCAGGGGG AATCTGAGGC CCAGAGAATT  
46151 GACTTTTCTT TTACAACAAA TGAGAGGGGG AATGCAGTAT CTTTGCCTCC  
46201 AGTGCTCCTG GTTCTCATGC TGCATGAAAC CTCTGAGGTC TCATTTTCTT  
46251 TCATTCTGGG ATGGGGATAA GAATATCTAA TAAGAATGGT TTAAGAATCA  
46301 AGCAATATCA GGTATGTGAT AATGTCTGGT AACTGGAAT AACCTATTGG  
46351 AACATAGTAG TTGTTTACAA AATATTTTTA AAATTTGTT ATACTTATGG  
46401 TCAACACTTT TTATATTTGT CTGTAGATTT CTGTACAAAA AGATTCTGAC  
46451 ACTGTTTTAA GCCAGCATTC CTTCAGAATG TACCCAAATC TCAAAATTTA  
46501 TTTAGGGGCA AAGCTAATGC TTTAAAGAAA AAGGAGAGGG GATTGGTGTG  
46551 TGTTTTTCTT TAGGAACAGT AGTAACTTGA CTTTATAGAG ACTTGAATAA  
46601 GCATTTATTT TTTCTTTTGT CCTATTTTAT TGTGAAGTTT ATTTATTTAA  
46651 AATAAAATGG ATTTCTCTGG AATTTAGTTT CTGCAATTT GAGGAGTTTC  
46701 CAAAGTCAAC CTTCAGGTTT GATACTTCTC TAGAAAGACT CACATAACTC  
46751 ACTGAAAGCT TATTACCCCT GGTATGGTT TATTACGGGG AAAAGATGCG  
46801 GATGAAATC AGTCAAGTAA AGAAGCACAT AGGCGAGAGC TTCTGTTGTC  
46851 CTCTCCCTGT GGAGTCTCCA TGTCTTACTT TCCTGGCACT GTTATGTGGC  
46901 ACTAGGCATG GAATATTGCA GACCAACCAG GGAAGCTCAC CTGAGCCTTT  
46951 GGTGTGCAGA GTTCTTATG GGGCTGTTT TCATACTGGC CACATGGCTG  
47001 GCCTTCAGAA TTCAACCCGT TCTGTGAGTG TGTGTGTGTG TGTGTGTGTG  
47051 TGTGTGTGTG TGTTTAGTGG TAGTCACCCC TTTTATGTGA GCTGAAACAA  
47101 TCAGAAGAAT AGCTGATTTG TTTAATTATT TTTGGTGTAT TGGACTTAAT  
47151 CAGTTTTTAT CTGTAGGTGG TCATAAGGTA CAGTATTTTT AAGTGACTAC  
47201 CACATCTGTA GTATAAGCCA AGTAATTTAT CAGTACTCAC AGGATGGGTA

FIGURE 3, page 15 of 42

47251 CATGTTGTAA TGAATTTATT GCCTAGAGAG GGCCTCAAAA TATGCCAAAG  
 47301 AGGGTGCAAT TTTTATTTTT GGTTCAGGC TGTATGCATT CCAGTGTGG  
 47351 TAGCCCTGAT ATACACAATA TCCAAACCAT TTCAGACCA TTTACAGTTC  
 47401 ATGCTCTGTAC TACTTCTTGA GGAGAGGGAG TAACATATTA CTTTAAATTA  
 47451 TAGTAATAA TATACATACA TTAAATTATA TGTAATAATA TAATATTATT  
 47501 ATTTGCAGTA TACTTTTTTA TTTCCCTTTA ACTGAGCTTG TTCATGTTTC  
 47551 AAAGGGTGT CCATTGCCTG ATACATAATT TAGTTAATAT TATCTTATGA  
 47601 AGGTTGTTC TAATTTTAACT ACTCTTCTTG TCTTCTCTCT CTGCTTTCTC  
 47651 AACTGAAGA TACCAATTAT TCTTAGTTTT AGAGTCAGAG ACAGGCCTCT  
 47701 AAAATCATGG CAATACTCCC TCTCATCATT ATATATATTT TTCAACCTTT  
 47751 CTATATTTTA TTTTCAAATA TATCTTCTTG CAGTTAGAAA CGGTATTGAA  
 47801 AAAGATTGTG TGGTTGTTCT AGAAAAAGTA ATAGTAATAT GCCACCAGCA  
 47851 TTTTATATCA TTCTGCTTTT ATTTTtaggt TCACGGTTCA AAATCAGACA  
 47901 AAATGAACAT ATTTGGTGGC TTTCGACAGA TGGTAAAAGA AGGAGGTATC  
 47951 CGCTCGCTTT GGAGGGGAAA TGGTACAAAC GTCATCAAAA TTGCTCCTGA  
 48001 GACAGCTGTT AAATTCTGGG CATATGAACA GGTAATTGTT ATCACCCGTG  
 48051 GAATTTATTA ACAAAGAGGA GTTAGTAAAC GGATTCATA AATGTTAATG  
 48101 TATAATGCTT TTGGGATTCT TGTTTTAATA CATGATAATC TTTCACATAT  
 48151 ACCCATAGG GAGGATCACT TATAGGAGAT TAGACTAAAT AAAATCAGAG  
 48201 ATTTCTCATG ACCAAGTTAT GGGATTCTTA ATTCATCATA TTATTTATAA  
 48251 AGTTTTTTTT TTCTAAGTAG TTCTTAAAGG AAGGGTAGAA TTTTAGTTTA  
 48301 TTCATTCTGA ATCCTGAGCA GAAGCAGCAC ACTAACATAA GTTTTATGAA  
 48351 AGTGTACAA TCTAACCTCT GGAAGGAAAA CTATAAGTTG AAGTCCTTTG  
 48401 TGTAATTTGA CGTTGCTGTA AAATTGAGCT GAGTTTGGAG TGACACCTCC  
 48451 ATGAAGCAG GGGCGTGGCT TCTTCCCCAT GTACTCCAGC ACCTAGACAG  
 48501 AGCTTGCCAT GTGATAAGTT TCAAGCGAGT GTTGAATGAG TCAATGAATG  
 48551 AACAAATGCA TTTACCTCTG AATCACTTCT CTGTCGGCTT TTGTTAACTT  
 48601 GGATTATTTG AGCTATTGCT TCAGCCTAAC TCAATGTAAA GGGGAAATAC  
 48651 AGAGGTAAGT TTTAGAGTTT GGGTCTCTT TATGGTCATT AGCAGAACTG  
 48701 TCTAGTTGAG CAGCCACAGA TTATGTTTTT CATTATTTAT TCCATCATTG  
 48751 TTTATCAAGG ACTGTAAGGG CCTTGAAATT CAACTCCCC CCCCATAGTT  
 48801 TTTGTATTAT TCCATGTAGA TTTTAGATTA TTCTGGAGAG TGTTTTGTTC  
 48851 TTAGCAACA GAATACTCTT GAGAAGATTA CGAAGTCCAG TGGTATCCTT  
 48901 TTCTTTGCCT AGGAAATAGA GAAGCAAAAA AAAAAAAAAA AAAAAATTAA  
 48951 AGAAAATCTA GTCTCCAGGA TTTTAATTAG AACCTATCCT TGGGAAGGCT  
 49001 ATTTTCCTTA TATGAAGGTT TGAAGATTCA AATCATGATT ATTAAGGGCT  
 49051 AATGTTTGAG ATACCCTTAG GTTATTCTGA CCACATACTT GGATTTTATG  
 49101 ATAGGAAAGC CACAGCCTAA AATAAATAAA TACTCAATGC AGTTATTTCA  
 49151 GTATGCAAGA AGTTTGGTAT TTTTGAAAAA GTCCATGGGT ATTGCAAGCA  
 49201 AATATGCACA TTTTGCTTTA TGCCATTTGT CAGATTCTTA CCTTGGATAC  
 49251 CACCAACAGG CATCCTCTGC TTCTGTCCAC CCAAGCTCCT TCCTGAGACC  
 49301 TCTTTATAGT ATTGTGATTT CTGCACACTA ACTTCTTAG ACATGAAGAG  
 49351 AAAGCTGTCT ACACAGTGTG GTGTAGTTTT CTATGGGCT CTGGACCTAT  
 49401 GGTGCTGTTT TCTCTCCTCC TGCTGAAGGT CCATTCATCC CTCGGGGCTC  
 49451 TCTAAAAGCC ACCTTCTGT GACAAGCATA TACTAAGCAT CTCAATCAAA  
 49501 GCCAGTTTCT CCCCTGTCCA GCCTCCCTCG AGTGCTGAAT TGCAGAATAT  
 49551 CCCATTTTTT ATTGGATGAT GGAAAACCCA TTGTTTTCCC AGTGGATTGT  
 49601 AAATTACTTC GGGGTAAATA GGCTGTATAT ATTCTCAAA TTTCCAGAGT  
 49651 ATGTAAC TAGTACTTTA GGAACTAAG AAAAAGATCT TTTCAACCTG GTATGTAGCT  
 49751 CTGTCAAACA CATCATCAGT ATGGGGTAAA CCTGTGTTCT CTGTGGGTTG  
 49801 TCATTACCAT AGTAGGTCA TTGTATCATT GACAGTGTA TAGTGTGGGG  
 49851 TAGTGTCTTT GTGGTTTCAG CTGCCACTCT GTACTGACTG CTTTCCACTC  
 49901 CAACATCTTC CTCTTTATCT CAACACTGTA GGTCTACCTG TGTACTGTGT  
 49951 GTTTCAGCAT CTCTGCTTGC ATGACCCAGG AGTGCTCCC ACTCAATATG  
 50001 GCCACCATGC ATGGTCATCT TTCTGCTACT CCCTGTCTCC TGACCCTGCT  
 50051 CCAGCAACAC AGACAGACAC CCTTCTCTT TCTATATGTC ATATGGTGGG  
 50101 GAATGCCCTT TAGTACTTAC TCAGGAGTTA GTTCTCTGG GAAGCCTTCT  
 50151 GTTCTAGTTT CTTTGTGTTA CAGCACTTTC ACATTGAATT CTGACGTTCT  
 50201 CTGTAATTAT CTGCTTTGTG AGACTGTGAG CTTCTTAGG CAGTAGCTAC  
 50251 TTGTATTCTT AGCACCTTGC CCAGTGCCAG GAAACCTTA TTAAGTAAAT  
 50301 GAAAGACAG AACTGACAGA CTGGAATTAG AGCTCAAGCT TGCCCTCAATC  
 50351 TCAAGCCATT AAGATGAAGG GGAGCCGGGC GTGGTGGCTC ACGCCTCTAA

FIGURE 3, page 16 of 42

50401 TCCCAGCACT TTAGGAGGTA GTTTGCTTGA GCCCAGGAGT TCAAGACCAG  
50451 CCTGGGCAAC GTGGCAAAAC CCCATTCTA CAAAAAATAT AAAAATTAGT  
50501 TGGACGTGGG GGTGTGTGCC TGTACTCAGG ATGCTGAGGT GGGAGGATCA  
50551 CTTGAGCTCG AGAGGCAGAG GTTGCACTGA GCTGGGATCA CACCATTGCA  
50601 ATCTAGCCTG GGTGATAGAA TGAGACCTTG TCTCAAAAAA AAAATAAATA  
50651 AATAAATAAA GGGGAAGATA AGGATTGGAA ACAGAAGGAG CAGCATGTGG  
50701 ACAGAAATGT AGGCACAAGA AGGCATCACT CACTGAAGAG ACTGAAAGTG  
50751 GTTCACTGTG CCTCAAGACT GGTGGAGTGT GTTTCGGGAA AGATAATGAT  
50801 GAAAGAGCTG GACAGATAAA CAGGGGCCAA ATGTAATAGG AGTCTGGATT  
50851 TTATTCTGAA TATGGTAGGG GCTATTGTAG CATCTTATAT AGGGAAGTGA  
50901 AATGAGTACA TTCACATTTA AGGAATATCA ACCTGAAAAA AGAGTGGAGA  
50951 CATGTGTTGGG GGAGATGAG GTAGACTAGA GGCAGGGAGA ATATTTAAAT  
51001 AATTGAGGTA AGAAATGATG AACACCAGTA TAAGGTGATG TCTTTAAGGA  
51051 ATGGAGAAGG GAATGAAC TGAAATATTT TGGAAGTAGA ATCAACAGAA  
51101 CTCCTGACT GACTGGATAT GGAGGTGAGA AAGAGAAGAG TCAAGAATGA  
51151 TATTCTAATT TCTAACTTGA GTGACTGCAT TCAAAGAGAA TACAATATCA  
51201 GGTTCATT TGTGCATGCT GAGTTTGAGA TGTGTGGGAC ATGTACAGGG  
51251 AGCTGTCCAG TAAGCAATTG GGTATATCAG CTAGCCATTA AGAGAGAGAT  
51301 CTTTGATAGA GAGGTTGTTG CTGAGTTGAG CCATTGGAAT GGGCAGGATC  
51351 ACTCAAGAA AGCTTATAAA TGAGAAGAAT TCTAGGAATA AGTCCAAAGG  
51401 GAGAAGTAAA AGAAGAACT TGCAAAGGAC ACTGAGAAGA AATAGCTCGA  
51451 GGGATGGGAG AAAATCCAGA GAGAGGGATG GCATAGGAGT CAGTGAAGG  
51501 AAACGTTTC ATGGGGGTCA GTACTACTGG GTAGTGAATA TAATAAGAAT  
51551 ATCTTTTAGG ATTTCTCAAC CCAGAGATAG GTAAGCTTAG TATAAATGCT  
51601 TCTGTGAAGT AATGAAATGA GAAACCATGC TGAAATGAGC TTAAAGTGAA  
51651 TGGGAGGTGA AGAACTTGG ACAGTAGAGA CACATTTTGA GGGAGTTTGA  
51701 CAGTGAAGAG AAGGAACTA GAAGAGGGAG AGGGTGATAG ATAAGAAAGA  
51751 TGTGGGTGG AGGGGATTG TTTTTTTGTT TTTTGTGTTT TTTTCTGTTT  
51801 GTATGTTTGT TTGTTTTTGA GATGGAGTCT CACTTTATCA CCCAGGCTGG  
51851 AGTAAAGTGG TGCAATCTCA TCTCACTGCA ACCTCTGCCT CCTAGGTTCA  
51901 AGTGATTCTT CTGCCTCAAC CTCCTGAGTA GTNNNNNNNN NNNNNNNNNN  
51951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN  
52001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN CAGCCTCCCG AAATGCTGGG  
52051 ATTGCAGGAG TGAGCCCCC GTGCCTGGCC TGGAGGGAGG ATTTTGATTT  
52101 GACTTTAATG TGCCTGTTGC TGAAGGAAGC ATGTCAATAC AAATAAAGAA  
52151 GTTGAAACA TAGGTAAGAG AGGTTGATTA ACCCGGTAGG TGTTTCAAGG  
52201 GAGTTTGTGT GTAGGGAAG GGAGTGGGAG ATGGAAAGG GCTGGGGGAG  
52251 ACAGGTCTA TCCAGAGACT GTTAAAGGA TTAGTCTTTG ATTACAAGAA  
52301 GAACCTCTT TATACGTGTT TGGGAAGAAA AAATATGTGA GTAGCTATGG  
52351 ATAATTTTGC AGGAGGTGG CAGAATACCA AGATATTCTG CCTGGTGGCC  
52401 TCTCTACTCT TCCTTGAGCT CCTGAGAAAG GATGTGATCT GAGAATGAGG  
52451 GAGGAAGTGG TATTGGAAGC TGGAGGAGAA TGGAGAAGAT CAAAATGGTT  
52501 AGTCTAACAA ATGGGAGAGA ACTGAGATAG ACAAAGGAT TTCAGGTGG  
52551 TTTTGAGGGC TCAGTTAAGT CTCCTTTAGG AAGGTTCACT TCTGTAGCCT  
52601 TGGCAAGTTA CTTAAAGTCT CTGTGACTAT TACCTCATCT CTAAGATGGG  
52651 GACTAAGCTT GGTGACATAG TTTTACATAC CAGGCACAGT GCCTGACTTT  
52701 TTGGCTCTGT CCTGAAGTCT TCCCTTTGTA TATGGTATGT TTCGGGGAAT  
52751 AGGAGCCTCA AGCACTTATC CTTTAAATAT TTATCCTCCA TCAGTCACTA  
52801 AACGTTTACT CTGTACTTTT GATAGGTGCT GTGGGGGTCC AGGGTATAAA  
52851 AGGTACCTTC AAAGTTACTG TTAAAGTGCA GGAAGGTTT TAAGCAAATT  
52901 ATGTTTAATG ATTTTGACAA TCTGACATGC AGGAAAATTA ATAGGGCTTA  
52951 TGCAGAAGAG GAGTTTTATG TAACACTCTG TAGTTCAGGA AACAGAGCCC  
53001 TTGGAAGCAG TGATCTCTCT GGGGAGGAAT GTCTGGTATT TGGGAATCTC  
53051 ATGAAATGAT AATATACTTA ATTTTATCA TGAGCAGCAA AACACAGATT  
53101 TGCTAGGAGA AAGTCATCGT ATGTTGTTGC ATTGGGCACT TTAGATCCCA  
53151 GGGAACAGAA ACTGGCTGGC ACAGGAATGG GCATCACTGT GGGGATGGAT  
53201 CATGTAGGGG AAGGATCCCT GGAGAAGTCC AGGAGGTGAG ACTTCCCCCT  
53251 TCCCTTCTCC ATGCATGAGT CCACTTCTCT CTGTTGACTT TCCCCTTGTC  
53301 CCTCTGGTGA CAGCAGCTGC TTACCTCTGG AGACCCCTC ACATTTCTGA  
53351 GAGAAGGAAT CTGGCTTGCC TGGCTAATTC CCATGGTCTA TGTTTGGGCA  
53401 GAATGTCTTA GCAAGTTGTG TAAAGATAGT GTATTATAT ATTAATAATA  
53451 ATAATAACAT CTAAGTGAACA TTTGCTAGGT GTTCAGACCT GCACTAACCG  
53501 TGTTACAAGT ATTATTTTTT TGTAATCCTT TCCATAACCC TGTGAGGTAA

FIGURE 3, page 17 of 42



53551 GTACTGTTAT CACAGACAAG GAAACCACAA TGTGGACCTG TTCATGAACT  
53601 TGCTCGAGGC CACGTGGCTC TGGAGTTCCA GCTCAGGTCT GCCTGACTCT  
53651 CAATCCCATG ATATTAATAT ACTGGCCAGT CACTATTTTG GCTGTATTGG  
53701 GGTCATATTT ATACCCTTGG TCCAGTTAGC TATGTTGGGT CACTTTAGTA  
53751 CTGATAGCCA GGGAGATGCT GGGCTTGATA GGTTAGTATA ATTCTATGTA  
53801 TTACCTACAA AAACGTGTTTT TATAAATTGT TTTGTTAACA TTTGTTTGTC  
53851 ACCTATTTAT TCATTTTATT TGCACCTGGT AAAATAAACT CATCTTTTAA  
53901 AAACGTGGG GAAAATATCC AAACATTGTG AAAACTTGAT TAACCTTGTA  
53951 TTTTCTGTAC ACCTGGGGAG GGATGCTGTT ATGCTGTTTC AGCAAAGGAG  
54001 CAACTTGGTC CAATCTGGGA GACATCTGTG TTTTGTGGAA ATCTGACTTG  
54051 AAAACCACTG TCCAGTCACT GCGTGTATTA GCATTTAGGC CTTGCTCTTC  
54101 TGCTATGTAT TATTAATGTA GTGTATACAT TTCGAGACAT ATCATCACAT  
54151 TTGTCAATTT ATTGATTTCT AGGAGCTGAT TTGTATTCTA GGATTGTCTA  
54201 GTTGGCTTGG GCTGCCATAA AATACCACAG TGTGTGTGGA ATCAACAACG  
54251 GAAATTTATT TCTAACAGTT TCAGAGGCGG GAAAGCCTAA GATCAAGGGC  
54301 CAAGCCAGTT TGATTTCTAG TGAGCGTTCT CTTCTCAGCT TGTAGACAGC  
54351 TGGTATGTGC TCACATGGTC TTTTCTTGGT GCACATGTGA AGGGGGAGAG  
54401 AGAGAGTGGG CTCTCTGGTG TCTGCTCTTA CAAGAACACT GATCCTGTCA  
54451 TGAGGGCTCC ATCCTCATGA CTCATAACC CTAATTACCT CCAGAAGCCT  
54501 CATCTCCTAA TACCATCACA TGGGAGGTTA CAGCTTCAAC ATATGAATTT  
54551 GGTGGGGGTG CAGCTCAGTC CACAGCAGGT AGTAATGTGC ATTTTAAAC  
54601 TTGTTTATAC AGTACAAGAA GTTACTTACT GAAGAAGGAC AAAAAATAGG  
54651 AACATTTGAG AGATTTATTT CTGGTTCCAT GGCTGGAGCA ACTGCACAGA  
54701 CTTTTATATA TCCATGGAG GTGAGTACCA TTGTCAAGTC TGAATGTGTG  
54751 ATGGTGTTCG TGTTGGTTGT CTATTGCTCT CTAACAAGTT ATCCCAAAAT  
54801 TAACAGTTTA AAACAAGCAT TTATCATCGC ACAGTTTCTC TGGGTGAGGA  
54851 ATCTGGAAGC AGCTTAGCTG GGTGCCTCTG GCTCAGGGTT TTTACAGCC  
54901 CACAGTCAAG ATGGTAGTCA GAGCTTGGAA TCAGCTGGAG GCGGATTCCA  
54951 AGCTCACTCA TGTTGCTGCC AGGCTCACT GGCTATTGGC TGGAAACATC  
55001 AGTTCCCTAT CACGTGAGCC TTTCTGTAGG CTGCTGAGT ATCCTCAAAA  
55051 CACAGTAGCT GGCTTCCCTA GAGTCAGTGG TCCAACAGAG AGAGAGAGAG  
55101 AGAGTGCCTA AGATGAAAGC TGGTATCTTT TGCTCTTCT GCTGTATTCC  
55151 ATTGATCACA CAGACCAACC CTGGTAGAGT GTAGGAGGGG CTGGTATAAT  
55201 GGTGTTAATA ACCGGAGACA AATATCACTG GGGGTCACTT TAGAGGCTGG  
55251 CTGCCACTTT AGAGGCTGGC TGCCATTCTT GTCCAAAGAG TTTCTGTACC  
55301 ATAAATTTAA TAATGGAATC TCAGGATTTG ATTATATGGT GATTATCCTA  
55351 ATTAGACATC CTTTCATTAG TGCATAGGTT GGCAAAACAC AGACCTACGG  
55401 ACTGTTTCAT ACAGCCCTTG ACCTAAGAAT GCCTTTTACA TTTTAAAAAA  
55451 GTGGGCAACA CAGGAAAAAG TGAGAAAGAT CTAAATCGA CACCCTAAGA  
55501 TCACAATTAA AAGAACTAGA GAAGCAAGAG CAAACAAAT CAAAAGATAG  
55551 CGGAAGACAA GAAGTAGCTA AGGTCAGAGC AGAACTGAAG GAGATAGAGA  
55601 CACGAAAAAC CCTTCCAAA ATCATTGAAT CCAGGAGCTG TTTTATGAA  
55651 AAGTTTAAACA AAATAGACAA CTAGCCAGAA TAATAAAGAA GAAACCAGAG  
55701 GAGAATCAAA TAGCCCCAAT AAAAAATGAT AAAGGGGATA TCACCACCAA  
55751 TCCCACAGAA ATACAACTA CCATCAGGGA ATACTATAAA CACCTCTATG  
55801 CAAATAAACT AGAAAATCTA GAAGAAATGG ATAAATTCCT GGACACATAC  
55851 ACGCTCCCAA GACTAAATCA GGAAGAAGCT GAATCCCTGT ATAGACCAAT  
55901 AACATGTTCT GAAATTGAGG CAGTAATTAA TAGCCTACCA ACCAAAAAAA  
55951 ACCAGGAGC AGACAGATTC ATAGCCGAAT TCTACCAGAG GTACAAAGAG  
56001 GAGCTGATGC CATTCCTTCT GAAATTATTC AAACAATAGA AAAAGAGAGA  
56051 TTCTTCCCTA ACTCATTTTA TGAGGGCAGC ATCATTCTGA TACTAAAACC  
56101 TGGCAGAGAC ACAACCAAAA TAGAAAATTT CAGGCCAATA TCCCTGATGA  
56151 ACATCAATGT GAAAATCCTC AATAAAATAC TGGCAAACCTG AATGCAGCAG  
56201 GACATCCAAA AGTTTATCCA CCATGATCAA GTTGGCTTCA TCCCTGGGAT  
56251 GCAAGGCTGT TCAACATATG CAAATCAATA TAACGGAATT CATCAATAAA  
56301 CAGAACCAAGT GACAAAAACC GCATGATTAT CTCAATAGAT GCAGAAAAGG  
56351 CCTTCGATAA AATTCAACAC CACTTCATGT TAAAACTCT CACTAACTA  
56401 GTTATTGATG GAATGTATAA CAAAATAATA AGAGCTGTTT ATGACAAAAC  
56451 CACAGCCAAT ATCATCTGA ATGGGCAAAA GCTGGAAGCA TTCCCTTTGA  
56501 AAACCGGCAC AAGACAAGGA TGTCTCTGT CAGCACTCCT ATTCAACGTA  
56551 GTATTGGAAG TTCTGGCCAA GGCAATCAGG CAGGAGAAAG AAATAAAGCG  
56601 TATTGAGATA GGAAAAGAGG AAGTCAAATT GTCTCTGTTT GCAGTTGACA  
56651 TGATTGTATA TTTAGAAAAC CTCCTTGTCT CAGCCCCAAA TCTCCTTAAG

FIGURE 3, page 18 of 42



56701 CTGATAAGCA ACTTAAAGCA AAGTCTCAGG GTACAAAATC AATGTGCAAA  
56751 AATCACTAGC ATTCCCTATTA ACCAATAATA CACAAACAGA GAGCCAAATC  
56801 ACGAGTGAAC TCCCATCCAC AATTGCTACA AAGAGAATAA AATACCTCGG  
56851 AATACAACCTT ACAAGGGATG TGAAGGACCT GTTCAAGGAG AACTACAAAC  
56901 CACTCCTCAA GGAAATAAGA GAGGACACAA ACAAATGGAA AAACATTTC  
56951 TGTGCATGGA TAGGAAGAAT CAATATCATA TCATAGGAAG AATCAGTGGC  
57001 CATACTGCCC AAAGTAATTT ATAGATTCAA TGATATCCCC ATCAAGCTAA  
57051 CATTGAATTT CTTCACAGAA ATAGAAAAAA CTACCTTAAA TTTTCATATGA  
57101 AACTAAAAAA GAGCCTGTAT AGCCAAGACA ATCCTAAGCA AAATGAACGA  
57151 AGCTGGAGGC ATCACGCTAC CTGACTTCAA ACATACTACA AGGCTACAGT  
57201 AACCAAAACA GCATGGTACT GGTACCAAAC AGATATATAG ACCAATGGAA  
57251 CAGAACAGAG GCCTCAGAAA TAACACCACA CGTCTACAA CATCTGATCT  
57301 TTGACAAAAA CAAGCAATGG GGAAAGGATT CCTTATTTAA TGTATGGTGT  
57351 TGGGAAAACCT GGCTAGCCAT ATGCAGAAAA CTGAAACTGG ACCCCTTCCT  
57401 TACACCTTAT AAAAAAAAAA TTAACCTAAG ATAGATTAAA GTCTTAAACA  
57451 TAGACTTAAA CTATAAAATC CCTAGAAAAA AACCAGGGCA ATACCATTCA  
57501 GGACACAGGC ATGGACAAAG ACTTCATGAC TGAATCACA AAGCAATGGC  
57551 AACAAAAGCC AAAATTGACA AATGGGATCT AATTAAACTA AAGATCTTCT  
57601 GCACAGCAAA AGAAACTATC ATCAGAGTGA ACCGGCAACC TACAGAATGG  
57651 GAGAAAAATT TTGCAATCTA TCCATCTGAC AAAGGGCTAA TATCCAGAAT  
57701 CTATAAGGAA CTTAAGCAAA TTTACAAGAA AAAAAACCC ACCAAAAAGT  
57751 GGGTGACGGA TATGAACAGA CACTTCTCAT AAGAAGACAT TTATGCAGCC  
57801 AACAAACGTG AGAAAAGGCT CATCATCCCT GGTGTGTTAGA GAAATGCAAA  
57851 TCAAAACCCC AATGGCATA CACTCTCAGC CAGTTAGTTA AAAAGTCAGG  
57901 AAACAACAGA TGCTGGCAAA TATGTGGAGA AATAGGAATG CTTTTACACT  
57951 GTTGGTGGGA GTGTAATTA GTTCAAGCAT TGTGGAAGAC AGTGTGGCAA  
58001 TTCCTCAAGG ATCTAGAACC AGAAATACCG TTTGACCCAG CAATCCCATT  
58051 GCTGGTTATA TACTCAAAGG ATTATAGATT TTTCTACTAT AAAGACACAT  
58101 GCACACGTAT ATTTATTGCA GCACGTGTTCA CAATAGCAAA GACTTGGAAC  
58151 CAACCCAAAT GCCCATCAGT GATAGACTAG ATAAACAAAA TATGGCACAT  
58201 ATACACCATG GAATACTATG CAGCCATAAA CAAGGATGAG TTCATGTCCT  
58251 TTGTAGGGAC ATGGATGAAG CTGGAAGCCA TCATTCTCAG CAACCTAACA  
58301 CAGGAACAGA AAACCAACA CCACATGTTT TCACTCATAA GTTGGAGTTG  
58351 AACAATGAGA ATACATGGAC ACAGGGAGGG GAACATCACA CACTGGGGCC  
58401 TTTTGGGGGA TGAGGGGCTA GGGGAGGAAT AGCATTAGAA GAAATACCTA  
58451 ATGTAGGTGA CAGGTTGATG GGTGCAGCAA ACCACCATGG CACGTGTATA  
58501 CCTATGTAAC AAACCTGCAC GTTCTGCACA TGTATCCCAG AACTTAAAGT  
58551 ACAATTTTTA AAAAGTAGGC AAAAAACAAA GAAAGAAAAA GTAATATACA  
58601 ACCGAGACCT AATATTTTAG GCTTGCAACG ACAGATATTT TACTATTTAG  
58651 TCTTTACAGG AAAAGTTTTT CAACTACTGC TTTATAGCAA AAATAATATT  
58701 GTAGATGTGG AATTTATTGA TATAGCAGAG GGGTTTTTAG TAACTGATGA  
58751 CTTAAGCAAG ATAAATACAA TTTTCACCGA TATGTGGTAT GCATGCTAAT  
58801 ACAGCTTTTT TTAAGCATCT TAATATGATT GTTTATATTA CTCCACACAC  
58851 CTCTCAAGAA AACTTAATAC CCTATTTTTC CTCTCATATC CTCCCATATC  
58901 AGTTAATAGT ATCACCTTCC CAACTCCCCA CTGCCCCATC CTGTGTTCCA  
58951 AGCTAGAAGT ATTGGGGTTA TCCTTTATAC TACCATTTCCT CTCACCTTCC  
59001 AGATGCAGGT GGTCAACAGT CAGTTTGTGT AAGACATCAA TAGATTATCT  
59051 TGCTTCCATT TCCTTGGTCA CTTCTTTCAT CAGATCCTCC TTGCAGTAAA  
59101 CGGGTCTCTC TGGCTTTGGT CTTAGCCCCC CAATAGAGGT AATACATGAA  
59151 AGAGAAATGTA TCAACAAATT GTACAGTCTT TTGAGTGACA ATATGTGCTA  
59201 GGTATTTGTT CCATGTAAAA TTACTTCATT TGAATCCCAT GATGATAGAG  
59251 TTAATATGAA CAATCATATT TTGTTTTTTT TTATATCCAG GTTATGAAAA  
59301 CCAGGCTGGC TGTAGGCAAA ACTGGGCAGT ACTCTGGAAT ATATGATTGT  
59351 GCCAAGAAGA TTTTGAAACA TGAAGGCTTG GGAGCTTTTT ACAAGGCTA  
59401 TGTTCCCAAT TTATTAGGTA TCATACCTTA TGCAGGCATA GATCTTGCTG  
59451 TGTATGAGGT GAGTTTGTAG AAATCTTTTG AATTGGAAAA TGCAGTTAGA  
59501 TCTTGTTAGA ATTGGACTTT ATATGAAGAA GTAGATATAT ACCAGAAAAC  
59551 AGTGTGTGAC CAGAAGTAAA TTCAAGCATG TGTATTTTGA ACTTTCAAGT  
59601 AACTTGAGTG TGAATATGCA TGGGGTCACT TTTGTATTAG ATTTTCTTGG  
59651 GAATTGCTTT TGTTAATGAA GAGTAGACTC AAAGTTAGGT ATAGTTGTTC  
59701 ACCTTAAAAG GTGTTTCTAG AGATTTTTTC CTTTGTTTTG GATTTGCAAA  
59751 AATCTGACAT TAAGCCAAGT GACTAATGTG ACTAACATGA GTAATACAGT  
59801 TTCATTCTTT GTACGGAAGA ATACAAATCT TGGATCAACC CTGCAATCTA

FIGURE 3, page 19 of 42

59851 AATCATTTAA TAATTTATGA ATCTCACAAA CAATTATTGA GCACACACTA  
59901 TACAAACCAC TAGGTTAGAC ACTGGATCTG GGGATTCAAA GGACTCAATG  
59951 TGTGCCTTGA AGAAACTGAA GGTCTGGTGG GGGAGACAAA CGACTAAAAC  
60001 TCAGCGTGGT TATCTGTGCT GCGACAGACA TGAGCCAGGG TGCATGTTAG  
60051 GATGAGACCT AAGCTACAGC GTAGAGGAAG AGTGAATGT GTAATGAAAA  
60101 GAAGAGTCGA ATTTTTTTTT TAAAGAGCTT TATTGAGATT TAGTTCATAT  
60151 TCCTTACATT TCACTCATTT GAAGTGTACA AGCAAATGGT TTTTGGCTTC  
60201 TTACATAATT TTTAAAAATT ATTATAAAAT ATAAATTTTG CCATTTTACT  
60251 AATTTTAAGT GTACAATTCA GTGGCATTAA TTACATTCAC AATATTGTGC  
60301 AACCATCAAC ACTATTTCCA AATCCTTTTC CTCACTCCAA ACAGAAACAC  
60351 CTTAACCTTT AAGCAATAAC TTCCTACCCT CCGTAACTCA AACCTTTGGT  
60401 AACCTCTAAT AGCTTTTCTA TGTCTAGGAA TTTACCCATT CAAGATATCT  
60451 TATAAGTAGA ATCATACAGT ATTTTCTTTT TTGTGTCTGA TTTATTACTC  
60501 TTAGCATAAT GTCTCTAAGG TTTGTTTCATG TTGTAGCATG TATCAGAACT  
60551 TCATTTCTTT TCATGGCTGA GTAATATTCC GTTATGTGTA TATACCACAT  
60601 TTTGTTTAGT CTTTCATCTG TTGAAGAGCA TTTGGATTAT TTCTACTTTT  
60651 CCAACATTGT GAATAATGCT GCAGTGAACA TTGGCATCTG CGTATCTGTT  
60701 CGAGTCTATG CCTTCAATTC CTTTGGGTAT ATATCTCAGA ATGGAATTGC  
60751 TGAGCCATAT GGTCAATCTG TGTTTAGCTT TTAGGAACTA TGAGACTGTT  
60801 TTCCATAGTG GCTGCACTTA CATTCTCACC AGCAACATAC AAAGGTTCCA  
60851 GTTTTTCAC GTCCTTATTA ACACCTAATT TCCATTTTAA AAAAGCTTAT  
60901 TTTTATATAT GCCGTCCTCT TAGGTGTGAG GTGGTATGGT TCAGGACTTT  
60951 ACTTCTGTG CTGAGTTTTT TAAAAAATTG TGATTAAAAA CACATAACAT  
61001 AAAGTTTATG ATTTTAACCA TTTTAAATA TATAGTACAG TAAGTGTTAA  
61051 CTGTTTGTGG TTTGTTGTGC AACAGATCTC TAGAACTTTT TCACTTCTCA  
61101 AAACCTTAAAC TCTATAGTCA TTAACAACA GCTCCCAATT TCCCCTTCAC  
61151 CCCAGCGCTG TGTAACCTAC TTTCTCGTTT TATGAGTTTG ACTACATTAA  
61201 ATACCTTGTA TAAGTGAAAT CATGTGGTAT TTCTCTTTCC GTGACTGGCT  
61251 TATTTTCATG AACATAGTTT CCTCATGATT CATCCATATG ATAGCATACA  
61301 ACAGGACTTT TTTGTTTTTA AGGCTGAATA ATAATTTGTT GGGTATATAT  
61351 ATCACATTTT CTTTATTCAT CTGTTGATGG ACATTTGGAT TGTTTCTACA  
61401 TCTTGACTAT TGTGAATAGT GCTGCACTGA ACATGGTTGT GCAAATATCT  
61451 CTTCAAGATA CTGTTTTCAG TTCTTTTGA CATATACTCA GAAGTGGAAAT  
61501 TTCTGGGTCA AATGGTAATT CTATTTTAA GTTTTGGAG AACCTCCATG  
61551 TCATTTTCCA TAGTAACCTAG ACCTTTTGT TTTTAAACAT TTCTATCAAT  
61601 GTACACCAAG ATTCCAATTT CTCCATGTCC TCCCCAACAC CATTAAAGTG  
61651 GGTGGTGGTC TACTACTATT GCTGTGTTGC TGTTTATTCC TCCCCTCAGT  
61701 TCTGTAAGTG TTTGCTTCAT ATATTTAGGA GCTTAATATT AGGTCCATAT  
61751 GAAGTTATAA TTTCTCCTG GTAAAGTGAC CCATTTATCA TTATGTAATG  
61801 TCCATCTTTG TCTCTTGTGA CAGTTTGTGT CTTAAATCT ATTTTGTCTG  
61851 ATGTAATTAT GGCCACCCCT TTTCTCTTTG GGTCCCCTT TTTATGGAAT  
61901 ATCTTTTTC ATCCTTTCAC TTTACGCTTA TGTGTGTCCT TAGATCTAAA  
61951 GTGAGTCTCA TAGATAAGGT ATAGTTGATT CTGTATGTGT TATTCACTCA  
62001 GCAATTTATA TCTTTTAGTT AGGGGATTTA ATCCATTTAC ATTTAAAGCA  
62051 GTTACTGATA GGGAAAGACT TACTGTTGTC ATTTGGCTAG CTACCTTTTT  
62101 ATCTTTGTCC TGTGGCTTFT CTGTTTTCCT CTCTCTCTCT TCCTGGCTTC  
62151 TTCTGTGTTT TGTTGATTTT TTTTTTTTTT GTAGTGATAT GTTCTGATTC  
62201 CCTTCTCATT TCCCTTTGTG TGCATTCTAT AGATGCTATT TTTGTGGTTA  
62251 CCATTGCAAC TACATAAAGC ATACTAAAGT TATAGCAACT TATTTTAAGC  
62301 TGTTTACAAC TTAACCTCAG TGGTATATAA AACTCTATTT CTTTACATAT  
62351 TTCACCTCCT CCCACAAAAC TTTATGTCTT TTGATATTGT ATATCCTTAA  
62401 CATAGATTTA TAGTTACTTT TTATGCTTTT CTCTTTTAAA TTCTGTTTAA  
62451 ATTTTGTTTT TGAAATTTAG ATTTTCAAGT TATTTATATA CCTTCATTAC  
62501 AATACTATAG GATTTTATAA TATTCTAAAT ATTGACCTTT ACCATAGAGT  
62551 TTCATATTTT GTGGTTTTGT GTTGCTATTT ATCATCCTTT TGTTTCTCCT  
62601 TTTAGCCTTT CTGTAGGGC CGGTCTAGTG GTGATAAGCT GTATCAGCTT  
62651 TTGTTTGTCA GGGACAGTCT TAATTTCTCC TTTTTTGAAG GGCAGTTTGG  
62701 CCCATACAGT ATTTTGTGTT GGCAGTTTTC TTAAGTTTCA AAACATAGAA  
62751 TATAACATTC TTTTTCCTTC TAACCTGCAA GATTTCCATT GAGAAATGCA  
62801 CTCAATGGAT TTTTAAATCC ATTGAGATAA TTTTTTAATC CTGTAGGATT  
62851 TAAAAATTTT AGTCTTACAG GATTAAAAAA TAAAAAGTT AAACCTGTTA  
62901 TATAACATAT TAACATGTAT TTTATACTTA AAGTATCTTA TGTTTAAAAA  
62951 GTTGATTATC ATATATATTT TATACAGTTT CTCCTAATTA TTGCCCTCTA

FIGURE 3, page 20 of 42

```

63001 ATGAAATACA GGGACCTAGA GTAACAGGGA TAAAGTATGG CCTTTTGATC
63051 AGCACGCCTG GTTCTGAGTC CTTCTTAAAA AAACCTCTGGG CCTGGTGTGG
63101 TGGCTCATGC CTATAATCTC AGCACTTTGG GAGGCCGAGG CGGGCGGATC
63151 ACCTGAGGTC AGGAGTTTGA GATCAGCCTT GCCAGCATGG TGAAACCCTG
63201 TCTCTACTAA CAGTACAAAG ATTAGCTGGG CGTGGTGGTG GGTGCCTGTA
63251 ATCCAAGCTA CTCAGGAGGC TGAGGCAGAA GAATCGTTTG AACCTGGGAG
63301 GCAGAGATTG GGCCACTGCA CTACAGCCTG GGTGACAAGA GCGAGACTCC
63351 ATCTCAAAAA AACAAACAAA AACTCCGCTG AGATGAATTT TTCTCATTTT
63401 TAAAAATCAGA ATAATAGATT TATGTAAGAG TTTCTGTAAG GCTCAAATGA
63451 AATATATGTA ACGTGTAATA TGAGATACAA TTAGTAGAAT TATATTATTT
63501 TATTAATACT CACCTAAGA GGTGTTCTTT AGATCCTGCA GCGTTTGCTG
63551 CGCAGTTTCA GTTTGTTTAG AAGAATGTCA GTAACCGGTG CAAACCTCAT
63601 GTGTTCCGCA CCCCCAGTGG CCTCCACCT CTCCACAGAG TCACCGCCTC
63651 CTGCAGTGCC TGCTGCTTCT GCAAATGCGT GGCTCATCC TGCAGAAACG
63701 GGGCTTCTCA TGAGGTTGAG AATAGCTGTG AAAATGTTTA CGTTGAAGTT
63751 GTAGAGTTTG TTAATTATTT TCTTCTTTAT TTCTCTGGCA GCTCTGAAG
63801 TCCTATTGGC TGGATAATTT TGCAAAAGAT TCTGTAAACC CTGGAGTCAT
63851 GGTGTTGCTG GGTGCGGTG CTTTATCCAG CACCTGTGGT CAGCTGGCCA
63901 GCTACCCATT GGCTTTGGTG AGAACTCGCA TGCAGGCTCA AGGTGAATTT
63951 TTGATTACAG AACCACACCG ATAAAAGTGC TGCACCAGTA ATGTGCTTTT
64001 AGAACTCCAA GTTCTACTAA GATGCAGACT GTAGTTTTAA GACAGTATTT
64051 CTCACACCTTT TTTTCATTAT TGCCTCCTTA AGGAATCTTT TCAGAAATTC
64101 TTTTTCTAAA TGCTCCCTCG TCATGAAATT TTAATGCGAC AGAAGCATTG
64151 CATATGTACT GTATGCATAC ATATGCCTTA TAGATAAACA GAGTACTATT
64201 TTTTTTGACT GTGTTACATG CACGTTTTAA GATTATAAGC TTTAGTATCT
64251 GATGGATTTG GGTTCAGATC CTTGCCTCAG ACTTCTGGG GTTTTTAATG
64301 GGAATGAAAA TTGTACAGTG TTGTAAGAAT TACCAACAAT ATAAATAAAG
64351 CATCTGGGT TTGTTAAATT TTTGGTAAAT GGTGGTTGGA ATCATTTTTT
64401 AGTGTGCGT AGACCCTACA AGTTTTGAGC TGTGATTCCT CCTCACTGTG
64451 ACCTGTCTC CATTGTTGGC TTTGATTACA CTGTACCATC CTGGTTGTTT
64501 TGCCAGCCCA TTGATACTT TTACCATTG CTGGCTTTTA TTGCTATCCC
64551 CACTGTATTA AAGTATGCAT TCAAATGCCT TTCTTTTCTC TTTGATGCTT
64601 TCCCTGGTCA GTCTTATCCA TTGTTTTCTT AAGTAGTACA CCTGGGCAT
64651 CTACAGCTCT ATTCCCAACC TCCCTTCCAA GTGCCAGCCA CAGCAACCCC
64701 AGCCAAGCAG TCAGTAACCTA ATTGGCAAAT ACTCCCTGAG CCATTGTCCC
64751 ATTCTAGACA CTGCCAGATG CTAGGGGTAG AGCAGTCAAC AAGTCAGGTG
64801 TGGCCCCGCC AGTGTAGAGT AGAGAAGACG TTATGTCCAG CAAGTAAACA
64851 ACCTGGTTAA ACCAACTCCT CTTTTGTTAG GGGAGCACAG AGCAAGGAGC
64901 TATAACCTAA CTTGGGCGCT GCAGAATGCT GTCAGTGAAG CTGAGACTGG
64951 AAAGATGAGT GGGAGTTAGC TGGGCACAGG CCAGTGGAGT GGGAACAGAA
65001 AACATTCCAG TTGAGGGAAA GCATGTGTGA AGACACTGAG GCAGGCACCA
65051 ACATGGTGTA TTTAAGGAGC TGAGAGACAG TCATGGCTGT AGAGAAAAAC
65101 ACAAAGTAGT GAACTACACG TTTCTTGTGT ATTCTCTCAT TTCACCATCA
65151 TAACCATCTT GGGGATGGGA ATACTAACAT TATCCCCATT TTTCAGATGA
65201 GCAACTGGGG CAGAGAGAAAT TTAAGTAACT CCCACAAGAT TATACCTGTG
65251 GTAAATAGTG GGAAGTAAAT TCAGACACAT GCAGTCTGAT TCTAACCTC
65301 CTGTCTGCCA GCTCTGATCC AGAACTTTGC ATGACTGATA CGGCTGATAG
65351 ATTGTCTATG GCTGATAGAC TGTCAATTCT GACCTAAAAG TCTGATCATT
65401 TTACATCTGT TCAGACATCT TTGCAGCCTT TCGGTGTGAG TTCCAAAGTT
65451 GTTAGTGGGA ATTTCAAAGC CTTTAATAAT CTAGCCCCAC TTTGTTCACT
65501 CTCTGTGTAA TAACCACATA CAACAATTGG CTGCATCTCC ATAGCACATG
65551 GTACTCCTCC CGTTGTCTTG GTTGTGCCAG CAACACTGGT TTTCGCTTTC
65601 TCTTCTGCTG TGTTGAGGTC ATTTCCAAGG CCCAGGTCTT TGTGCTTTTT
65651 CCCAAGCTTC CCAGAGCTTC TTCCATACTC CCCTTACTTC CTGAGATTTA
65701 ACTGTTCTCT CTTCAGCGCT TGTCTAGTAA GAAGGAGGCA GCAGCAGCAC
65751 TGTGGGGTGG TGGAAAGTGT ACCAGCTTTG GAGTCAGACC ATTGGATCTC
65801 AGCCCTACCA TTTTCTACTT AGATTTTTTT AGGACAAATT TCTCCATCTT
65851 TCTAAGCCTC CAATTGCTCA CTTACAAAAT TGATATAACA TTTACCTTGC
65901 AAGATTGGTA TGGAAGGTAA TTAACCCAGT ATTTAGAACA TAGTAATTAA
65951 TAAATAACTA TTATTACCAT CATTACTATA GTTAGGACAC TCACTGTTAG
66001 GTGCTATACA AAGAGGATCA TAAAAGGGAT GTTGTCTTGG GCTTCTTGGA
66051 ATAAATGTTG TCCTTTTACT GTATTTTAGA ATATCATTTCT GGGTCATAAT
66101 TGTTTGTGTG CATAATAATG AAACATACTT GAATATTAAA TTACCCTCTT

```

FIGURE 3, page 21 of 42

66151 TTTTATTTT TTAGCCATGT TAGAAGGTTT CCCACAGCTG AATATGGTTG  
66201 GCCTCTTTCG ACGAATTATT TCCAAAGAAG GAATACCAGG ACTTTACAGA  
66251 GGCATCACCC CAAACTTCAT GAAGGTGCTC CCTGCTGTAG GCATCAGTTA  
66301 TGTGGTTTAT GAAAATATGA AGCAAACTTT AGGAGTAACC CAGAAATGAT  
66351 GTTGCAATTTT TTGCTTTAGC CTGATAATTG AAACTTTCAA CAATCTCTGG  
66401 AGTGACTTTT TCTCCTCGAA TTGAAACAAG TCTATGGCAA AAGAAGCTGC  
66451 ATTTTTTTTCA CAAAAGGGAA GATGGTAACA ATGGTCACTT CAAACTTTTG  
66501 GGCTAAATTA TATGTACACA GAAATGTTCA AAATCATAGT TTTAATGTGT  
66551 TTTGAAAAGG CCACACAATT ATACTTTATC TTTTCTTAAT AATCCTGCAA  
66601 ATCTCTGCCC TGAATCCGAA ATCTGAAAAT GTACTGGCTT GAACAAAATT  
66651 TGTTTTGTGT GTTAGAGTTA TAAATCATT AATCTTTATT CGGGTGGTTT  
66701 ACGTTTTATGC CAGTTCCTTT ATATTTAAAT TTCTTGTTTT ATATATTTTG  
66751 AATGTCTTTA TAGATTTCTT TAAATTTCTT TATAGAACCA TTAATAGAAA  
66801 ATCATTACAT TTAATAATA CCTTACAGCA AAAGCATCCA AATAAGTATA  
66851 GGGTTTTATGT CCTTATTTTT CTTTCAGCTG AATACGAATG AGCACAGTGG  
66901 TGAATTTCT GAAGGGAAGT GATGAAATTA TATTTATTTT AGTGGGCACT  
66951 TTTCCATTTT ACCACTGTAC CATTATTTGG TTCCTGGAGT TATACACTAA  
67001 TTTTCAGTAT ATTTAGTTTA AATTACCAAC ACAAGGCAAT TTATTTGAAA  
67051 GATTCCGTTT ATCCTGCCAT TGCTTTGAAA AGCAGCAGGA AACGAAATCC  
67101 TTTGACTTGT ATCAGCTTCT GCAGAGCATC TTTGTTTTCC TTTGTCCTTT  
67151 GTTTCCTACC TTTTGAATCA GATCCGTTT TAGTCAGGAA GACTTCTTGG  
67201 GACCATTCTT AGTAACCTGA AATTTCTTTT TTAATTGCAT GAAGTGGATT  
67251 GATCATGAGC AAATGATGTG CTTATTTCTC CCTCACTGTT GAATATCTTT  
67301 GAACTTGCTG TTTTCAATAT GGGCAGCACA AAGGTGAGAG ATACATATTA  
67351 ATAGTAGTAT GTATTACTCT TATACATTAG ATACCTATAT TTAAATGAAA  
67401 GGCCCAATTT GTAAACATAT ACATTCATAT TCTCTCTTGC CCCAAGTTTT  
67451 AGGAACATGT TAGGATATAG GAGACTTAAT TTATAATAAT GAGAGCATTT  
67501 TTTTATTTTA CTAAAGCCAT TTTTATAGTC AACTATCTTT TCTTATTTGT  
67551 GTGATTAGAA CTTAGAAAAA TATTTACTAG TTGAAGTTAT TATCAGTTTT  
67601 TAATTTAGTT CTTAAACTCA TTTCACTTCT AATAATTTCT GTTATAAATT  
67651 GCCAGCATTT TAATGAAAAT CTAATGATGT AATAGGCATT TTCTTTATTT  
67701 GAACCTACCT CTTTTATTTT CTGAACCAAA GAGAAAGATG GACTGGTGTT  
67751 TGTGAAACAT TTTTAAAAAT GTAGTTTCAT TTATATTAGT TATGTTTGAT  
67801 AAATGTCTCA GTATTTTAT AATATGATAA GCCTGGGATT CTACTTTTAG  
67851 GGTTATTTGT ACTTTTGAGT AATATATAAA GTGACAATAT TAAGGTACAT  
67901 GATCAGCTCT TTCTATTTTT ACTCGTAAAA ATTATGGAAA TGAATAATTT  
67951 TGCTAACAACT TTTGAAATTT CAACTTCTG GAAAATATGA AAATATTCAT  
68001 TGTTTCATTAT GAATTTAAAT TGTAAGGTAT GAATGTGATT TGTCTGTACA  
68051 TCTTGTATCT TTTCCAAAAA ATGATTCTGT ATCTTTTGGA AAAAAGCCGA  
68101 GAGTTGAAGA TAGTATATTT CTGGTAGTAC TGAATATTTA CTTACAGTTT  
68151 CTATCAAAAA TATATATTTG TTTCTAAAAT TACTTGTTTT CCAGTTTTTA  
68201 TTTTTTTTAG AGAAAATCT TAAGTCTCAG TTTCCTAATT GAAAAAATAA  
68251 AATTATAAAT AAAGCAAAAA TTGTATCCTA CAGCTTAGCT AGCTTAGATG  
68301 TTTGGCACCA GTTTGAATCA TGCTTTTAC AGCTGGCTCC ATGTAGTCTT  
68351 TCCAAACATT TTGGCCTTTC CTGAGCAGCC CTGTAGATA TTGTCTGTAT  
68401 GATGCATTTT GACACAAGGT GATATTTTTT GTGATATCAA AATTCCACAT  
68451 TTACCCATTA GAGTTACAGC CCTGGGGTTC ACAGTACCAA GGGGGACCCA  
68501 GAGCCTCAGG ATTGGCCAGG CTCATTTTGC CGTGGAGTAT CAGTTTGTCT  
68551 TGAAATTGTG GGAAAAAATT CTAAGTTGAA TTCACTGGTA AGTAATTTTT  
68601 TAAAATTTCA TAATGCAGAT TACATCCAAA ATTTGATTTA AAAATTAATA  
68651 CATAAGACTG CAGAGAAAAT CTGCATTTCA ACTCCAATAC TATCCAGACT  
68701 TCAGAAATAA CTTATCAGTT ATTTCTGTAA GCTTCTTGCT TACCTGGATA  
68751 CCTGACAGGT GAGATGGCTG TAGCAGACAC TGGCAGTTCC CTGCCCACAC  
68801 ACCTGTCCCT GTCCACAGCT GCACAAGGCA GCTCTGTGTG CAATTGCCAG  
68851 CATCTGCTCC TCTGTCTCA GGGAACTTTT GTTAGAAAAA TGCTGCCATA  
68901 TTTGTTTCTC ACCTATTAGT CTTGTCTCCC AGTCAAGAGA ATAAATTTAT  
68951 GCAAGCAGAG ATTGTACTTT ACAGTATTTT GTCTTTGAGC TTGGCATTAG  
69001 GTTGCAATTTG TAAAAATGTG GCATGGCTTC CTCATCCCCC AATAGGAACT  
69051 TTGCCAGCCC TTTTGTCTCT ATGGAACCTC CTTTTTTGAA AAGAGCACCA  
69101 AAGGAGTAAA AATACTGTGG AGGGAGCAAC CCTCCTTTGC CATATGCTCT  
69151 CATTGGGAGA CATGTGAGC AGTCTGAAGT CATTTAGGCC ACTCTCTGGG  
69201 AGAGCACATC CTATGATGTT CTCCCAGCCT AGCCCCCTCC ACTGTGCTCA  
69251 AGTCCAAGCT GACCAGCTTT CTGACCACAG TGTAAACAAA GATGATTGTC

FIGURE 3, page 22 of 42

69301 AGTGGGCCCC AGAATCCTAT ACCCAGA

**FEATURES:**

Start: 2132  
Exon: 2132-2314  
Intron: 2315-17055  
Exon: 17056-17182  
Intron: 17183-20983  
Exon: 20984-21071  
Intron: 21072-41719  
Exon: 41720-41831  
Intron: 41832-45391  
Exon: 45392-45550  
Intron: 45551-47878  
Exon: 47879-48031  
Intron: 48032-54612  
Exon: 54613-54720  
Intron: 54721-59290  
Exon: 59291-59458  
Intron: 59459-63791  
Exon: 63792-63942  
Intron: 63943-66164  
Exon: 66165-66346  
Stop: 66347

**CHROMOSOME MAP POSITION:**

Chromosome 1

**ALLELIC VARIANTS (SNPs):**

DNA

Position	Major	Minor	Domain
1722	G	C A	Beyond ORF(5')
1767	C	G A	Beyond ORF(5')
1840	C	G	Beyond ORF(5')
1857	T	G	Beyond ORF(5')
1945	G	T	Beyond ORF(5')
2007	A	C	Beyond ORF(5')
2769	C	G	Intron
3664	C	T	Intron
3827	G	A	Intron
4113	C	T	Intron
4337	A	G	Intron
4473	G	A	Intron
6455	T	G	Intron
6533	T	G A	Intron
6919	G	C	Intron
7305	G	A	Intron
7340	A	G	Intron
7466	A	G	Intron
7589	G	C	Intron
7810	A	C	Intron
9104	G	A	Intron
9503	A	T	Intron
9898	G	C	Intron
10196	T	C	Intron
12327	C	G A	Intron
13749	G	A	Intron
14150	T	C	Intron
14529	G	A	Intron
14653	G	A	Intron
15871	A	G	Intron
19244	G	A	Intron

19387	T	G	Intron
19447	C	G	Intron
20076	T	C	Intron
20492	T	-	Intron
20868	T	C	Intron
20941	T	C	Intron
21116	C	T	Intron
21701	G	A	Intron
21710	A	-	Intron
21826	C	T	Intron
21840	-	T	Intron
21841	-	C T	Intron
21843	-	C	Intron
22045	C	A T	Intron
22061	G	T	Intron
22348	-	A G	Intron
22682	A	G T	Intron
22783	-	T	Intron
23448	A	G	Intron
24960	G	A	Intron
24983	T	C	Intron
25390	T	C	Intron
26060	C	T	Intron
30245	C	G	Intron
33664	G	T	Intron
33883	C	A	Intron
34373	G	A	Intron
34558	G	T	Intron
43929	T	A	Intron
44309	T	- C	Intron
44997	T	G	Intron
46538	A	G	Intron
48153	T	C	Intron
48288	G	T	Intron
48412	G	A	Intron
48446	C	G	Intron
48456	G	C	Intron
48789	C	-	Intron
48859	G	C	Intron
49126	A	G	Intron
49378	T	G	Intron
49482	A	C	Intron
49741	G	A	Intron
49840	A	G	Intron
50102	G	A	Intron
50109	C	G T	Intron
50747	G	A	Intron
51272	G	A	Intron
52842	G	A	Intron
61837	A	G	Intron
62018	A	G	Intron
65562	A	G	Intron
65780	G	A	Intron
66092	G	A	Intron
66617	C	T	Beyond ORF(3')
66892	G	A	Beyond ORF(3')
67263	G	A	Beyond ORF(3')
67651	G	T	Beyond ORF(3')
67935	C	T	Beyond ORF(3')
69000	T	G	Beyond ORF(3')
69134	C	T	Beyond ORF(3')

FIGURE 3, page 24 of 42

Context:

DNA  
Position  
1722

TTGCCACGCAGATGGCTGTTGATCTTTTCTGCAACAAATCCAGGAGTTTCTCCTTTTTG  
TTTTATAATTGCTCCAATAGATGCTTTAGGATTTAACTCTCTGCTTTTAAAGCAGAATC  
GCCATCCCAGGTGTGCAACCACGAAAAAATTAGACATCCGTGAGAGACAATGCCCTCCAT  
GGCCAGTTTCCAGGCAGAGAGAAGCAGCTCTGGGCTGACCGCCAAGGCTCCGGCCCCGAG  
AGGGTCTTTAAGTGGAGTAACCAGTCTTCAAGACCCCGCTCCCAAGCCACCGACGCGCTG  
[G, C, A]  
CGCTGCAGCCCTGGACCTGCTGGGGGCTCTTCTCGGACCCGCATGCTGACAGCGGGAC  
TGGCAACTGGGCAGAGGTGACCCCGGTCCGCACAGCACCTCCCGAGACCCAGCTCCCA  
GCTCCCTCACTTCCGGCTCTCTGGAGGCGGGCCCGGCCAGTGCCGCGGAGGCCAGCGCG  
CGAGCTCCTCCCGAGCAGCGGGGACGGCCACACCTGCGCGCGCGCGGGCTCGGGTG  
GGGTCTCCGCTCCTGCGCCCTGCGCGCGCAGCCGCACCCCGACGCGCCCCAAACGCT

1767

AGTTTCTCCTTTTTGTTTTATAATTGCTCCAATAGATGCTTTAGGATTTAACTCTCTGCT  
TTTTAAAGCAGAATCGCCATCCCAGGTGTGCAACCACGAAAAAATTAGACATCCGTGAGA  
GACAATGCCCTCCATGGCCAGTTTCCAGGCAGAGAGAAGCAGCTCTGGGCTGACCGCCA  
AGGCTCCGGCCCGAGAGGGTCTTTAAGTGGAGTAACCAGTCTTCAAGACCCCGCTCCCA  
GCCACCGACGCGCTGACGCTGCAGCCCTGGACCTGCTGGGGGCTCTTCTCGGACCCCG  
[C, G, A]  
TGCTGACAGCGGACTGGCAACTGGGCAGAGGTGACCCCGGTCCGCACAGCACCTCCC  
GAGACCCAGCTCCCAGCTCCTCACTTCCGGCTCTCTGGAGGCGGGCCCGGCCAGTGCCG  
CCGAGGCCAGCGCGGCGAGCTCCTCCCCAGCAGCGGGGACGGCCACACCTGCGCGCC  
CGCGGGCTCGGGTGGGTCTCCGCTCCTGCGCCCTGCGCGCCGACGCCGACCCCCGAC  
GGCGCCCCAAACGCTGTTGCGCCGCGCGCCCCGCCAGCCCGGCTCGCGCTGGTCCCGG

1840

TCGCCATCCCAGGTGTGCAACCACGAAAAAATTAGACATCCGTGAGAGACAATGCCCTCC  
ATGGCCAGTTTTCCAGGCAGAGAGAAGCAGCTCTGGGCTGACCGCCAAGGCTCCGGCCCC  
AGAGGGTCTTTAAGTGGAGTAACCAGTCTTCAAGACCCCGCTCCCAAGCCACCGACGCG  
TGACGCTGCAGCCCTGGACCTGCTGGGGGCTCTTCTCGGACCCCGCATGCTGACAGCGG  
GACTGGCAACTGGGCAGAGGTGACCCCGGTCCGCACAGCACCTCCCGAGACCCAGCTC  
[C, G]  
CAGCTCCCTCACTTCCGGCTCTCTGGAGGCGGGCCCGGCCAGTGCCGCGGAGGCCAGCGC  
GGCGAGCTCCTCCCGAGCAGCGCGGGACGGCCACACCTGCGCGCGCGCGGGCTCGGG  
TGGGGTCTCCGCTCCTGCGCCCTGCGCGCCGACGCCGACCCCCGAGCGCCCCAAACG  
CTGTTGCGCCGCGCGCCCCGCCAGCCCGGCTCGCGCTGGTCCCGGTCTCGCCCCGAG  
CCCTCGATCTCCCGTGACTTCTCGGCCAGGCCGCTGCGCCTCTGGGACCATGTTGCGC

1857

CAACCACGAAAAAATTAGACATCCGTGAGAGACAATGCCCTCCATGGCCAGTTTCCAGG  
CAGAGAGAAGCAGCTCTGGGCTGACCGCCAAGGCTCCGGCCCGAGAGGGTCTTTAAGTGG  
AGTAACCAGTCTTCAAGACCCCGCTCCCAAGCCACCGACGCGCTGACGCTGCAGCCCTGG  
ACCTGCTGGGGGCTCTTCTCGGACCCCGCATGCTGACAGCGGGACTGGCAACTGGGCAG  
AGGTCGACCCCGGTCCGCACAGCACCTCCCGAGACCCAGCTCCAGCTCCTCACTTCC  
[T, G]  
GCTCTCTGGAGGCGGGCCCGGCCAGTGCCGCCGAGGCCAGCGGGCGAGCTCCTCCCCAG  
CAGCGGCGGGACGGCCACACCTGCGCGCCGCGCGGGCTCGGGTGGGGTCTCCGCTCCTG  
CGCCCTGCGCGCCGACGCCGACCCCGACGGCGCCCCAAACGCTGTTGCGCCGCGCGCC  
CCGCCAGCCCGGCTCGCGCTGGTCCCGGTCTCGCCCCGACGCCCTCGATCTCCCGTGA  
CTTCTCGGCCAGGCCGCTGCGCCTCTGGGACCATGTTGCGCTGGCTGCGGGACTTCGT

1945

CAAGGCTCCGGCCCGAGAGGGTCTTTAAGTGGAGTAACCAGTCTTCAAGACCCCGCTCCC  
AAGCCACCGACGCGCTGACGCTGCAGCCCTGGACCTGCTGGGGGCTCTTCTCGGACCC  
GCATGCTGACAGCGGGACTGGCAACTGGGCAGAGTTCGACCCCGGTCCGCACAGCACCT  
CCCGAGACCCAGCTCCAGCTCCCTCACTTCCGGCTCTCTGGAGGCGGGCCCGGCCAGTG  
CCGCCGAGGCCAGCGCGGAGCTCCTCCCCAGCAGCGCGGGACGGCCACACCTGCGC  
[G, T]  
CCGCGCGGGCTCGGGTGGGTCTCCGCTCCTGCGCCCTGCGCGCCGACGCCGACCCCCG  
ACGGCGCCCCAAACGCTGTTGCGCCGCGCGCCCCGCCAGCCCGGCTCGCGCTGGTCCC  
GGTCTCGCCCCGACCCCTCGATCTCCCGTGAATCTCTCGGCCAGGCCGCTGCGCCTCT  
GGGACCATGTTGCGCTGGTTCGCGGACTTCGTGCTGCCACCGCGGCTGCCAGGACGCG  
GAGCAGCCGACGCGCTACGAGACCTCTTCCAGGCACTGGACCGCAATGGGGACGGAGTG

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted May 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2007 GCCACCGACGCGCTGACGCTGCAGCCCTGGACCTGCTGGGGGCTCTTCTCGGACCCGG  
ATGCTGACAGCGGGACTGGCAACTGGGCAGAGGTGACCCCGGGTCCGCACAGCACCTCC  
CGAGACCCAGCTCCCAGCTCCCTCACTTCCGGCTCTCTGGAGGCGGGCCCGGCCAGTGCC  
GCCGAGGCCAGCGCGGCGAGCTCCTCCCCAGCAGCGGCGGGACGGCCACACCTGCGCGC  
CGCGCGGGCTCGGGTGGGGTCTCCGCTCCTGCGCCCTGCGCGCCGACGCCGACCCCGA  
[A, C]  
GGCGCCCCAAACGCTGTTGCGCCGCGCGCCCGCCAGCCCGGCTCGCGCTGGTCCCGG  
TCTCGCCCCGAGCCCTCGATCTCCCGTGACTTCTCGGCCAGGCCGCTGCGCCTCTGG  
GACCATGTTGCGCTGGCTGCGGGACTTCGTGCTGCCACCGCGGCTGCCAGGACGCGGA  
GCAGCCGACGCGCTACGAGACCTCTTCCAGGCACTGGACCGCAATGGGGACGGAGTGGT  
GGACATCGCGGAGCTGCAGGAGGGGCTCAGGAACCTGGGCATCCCTCTGGGCCAGGACGC

2769 TGGGGCCGCGACCCGGGACCCCGGTAACAGAAGTGGGTCTAATACGAAAGTCTACTGGT  
ATTTGTCCAGATAAAATGAGTGTGTTGGACACTCTGGCCACGGGCACTGTTAAATTTTT  
AAGACACTTTTGTCTGAATCCATCCCAGGTTCTTTGTTTTCTGTTTTAATACCTGCGA  
ACATGTAATCCGTTTTAGCTGTGCACTTCAGTGGGTCCCAAGTTTTGTATAAAGGCGCA  
CACATTGATCTCTTTTGAAGCTGCTTTGTTACAGCAGCTATGTGTATTGTCTACTGTTT  
[C, G]  
AAAACTGTTTGAACCAATCGCGTGTTCCTGTTGAGAAGGAATGGCGGC  
ATTCCATTGTTTAAACATTCCTAGGTTAATGCCCTAGGTACATAAATGATCTGAAGGG  
TTGACTTGACCTGCGACTGAGCAATTTCTCTGAGTCATCTTAACTGTGCCCTG  
AACTTCTGCCCCTTTAGTAGGGTGGAGATATGTGGAACCTTCCAACCTGTTGAAGCGT  
TCCTTGACACTGGCATTCTCTTATCCAAAGAGGGAAGTGATTAGGTTACTATGAGGGCC

3664 GCTGATTGTCCAGAAATGGCCAGTTGGAGTTCCCCACCATGTCCAATCATTGGCTGGA  
AGCAGCCCAGGAAAGGACGACCTTGCTGCAGTGCATCAGCAGATGCCAGGGTTAGAGGC  
TAGAGAGTGGAAGTCAACTGTGTTCTCACAGTAGGTGCCTTTGAAGGGAGATCTCAGTG  
GTACAACCTCCATGGTCCCTACATATACAAAAGCTCTTTGGAGTGCTCAATGATTTTTAA  
GATTGTAAAGGGATCCTGAGATCAAAAAGCTTGAGAATTGCTGCTGTATCACCATTTTTA  
[C, T]  
GTAACGTCATCATATTCTGTTATATGTTTGTGTCATAGTATATGTTACCAATTCTTTTTA  
AATCACCTTTTACTTTATTGATAGTTTAAAAACGATTGTAAGTGAAATTGCAATGGATGT  
CCTTTGTATTCTTTTCTCATTCTGGTCCAGTTACTTTCTGAGGATAAATTTGAGGAGT  
GGACATTGCTGAGTCTGAAGGTAACACACATTTTAACTGGGATACGTATGCCTTTTCG  
AAACCTTAGACCATTTTCACTCTTTTGAAGTACAGTGTGCTTCTCCACATCCTCGCT

3827 GAAGGGAGATCTCAGTGGTACAACCTCCATGGTCCCTACATATACAAAAGCTCTTTGGAG  
TGCTCAATGATTTTTAAGATTGTAAGGGATCCTGAGATCAAAAAGCTTGAGAATTGCTG  
CTGTATCACCATTTTACGTAACGTCATCATATTCTGTTATATGTTTGTGTCATAGTATA  
TGTTACCAATTCTTTTAAATCACCTTTTACTTTATTGATAGTTTAAAAACGATTGTAAG  
TGAAATTGCAATGGATGTCCTTTGTATTCTTTCTCATTCTGGTCCAGTTACTTTTCGTA  
[G, A]  
GATAAATTTTGAGGAGTGGACATTGCTGAGTCTGAAGGTAACACACATTTTAACTGGGA  
TACGTATTGCCTTTTCGAAACCTTAGACCCATTTTCACTCTTTGACTGACAGTGTTCG  
TTCTCCACATCCTCGCTCATTGAGGATCAGTCTTTGTAAAGTCTCCTATTCTGCAGGT  
GAAATTCCTTTTCACTTCTGTCTTAGTCCATTTAGTGTGCTATAGTGAATATCTGAG  
ACAGGTAATTTATAAGAAAAGACATTTATTAGCTCACAGTTCCGCAGGCTGGGAAGT

4113 CAGTTACTTTCTGAGGATAAATTTTGAAGGAGTGGACATTGCTGAGTCTGAAGGTAACACA  
CATTTTAACTGGGATACGTATTGCCTTTTCGAAACCTTAGACCCATTTTCACTCTTTTG  
ACTGACAGTGTCTGCTTCTCCACATCCTCGCTCATTGAGGATCAGTCTTTGTAAAGTC  
TCCTATTCTGCAGGTGAAATTCCTTTTCACTTCTGTCTTAGTCCATTTAGTGTGCTAT  
AGTGAATATCTGAGACAGGTAATTTATAAGAAAAGACATTTATTAGCTCACAGTTC  
[C, T]  
GCAGGCTGGGAAGTTTAAAGAGCGTGGTGCTGGCATCTGCTGGACTCCTGGGGAGGGCTT  
TCCTGCTGTGTCAACATGGTGGAAAGTCAAAGTGAAGTGGACATGTGTGAAGAAGCA  
AAATCCGAGGGGTGCTTGCCTTTATAGCAACCCAGCCTCGAGGGAAGTATCCATTACT  
GAGGGAAGTAAATTCAGTCTCATGAGAGAGAACTCACTCACTACTGCAAGAATGACACC  
AAGCCATTGAGGAGTCTGCCTCCGTAACCTGACACCTCCTGCTAGGTCCCTCCTCC

4337 CATTTAGTGTGCTATAGTGAATATCTGAGACAGGGTAATTTATAAGAAAAGACATTT  
ATTTAGCTCACAGTTCGCGAGGCTGGGAAGTTTAAAGAGCGTGGTGTGCTGGCATCTGCTGG

FIGURE 3, page 26 of 42



ACTCCTGGGGAGGGCTTTCCTGCTGTGTACAAACATGGTGGAAAGTCAAAGTGGAAGTGG  
 ACATGTGTGAAGAAGCAAAATCCGAGGGGTGTCTGGCTTTATAGCAACCCAGCCTCGAG  
 GGAAGTATCCATTACTGAGGGAACATAATTCAGTCTCATGAGAGAGAGAACTCACTCACT  
 [A, G]  
 CTGCAAGAATGACACCAAGCCATTTCATGAGGGATCTGCCTCCGTAACCCTGACACCTCCT  
 GCTAGGTCCTCCTCCCAACACGGCCACATCAGGGATCAGACTTCAACATGAGTTTTTGT  
 GGGGACAAACAAAACGTAGCACTTGCCTTTTGGTTCTATTACATCCTCCACAGG  
 ATTGCATTATGCCTACCCATTTGGTGAGGGCAGTCTCTTTAATTGGTTTACTGATTCAA  
 ATGCTACCTCCTCCAGAGACATCCTCACAGACACACCCAGAAATCATGTTTTACCAGTT

4473 T TCCTGCTGTGTACAAACATGGTGGAAAGTCAAAGTGGAAGTGGACATGTGTGAAGAAGC  
 AAAATCCGAGGGGTGTCTGGCTTTATAGCAACCCAGCCTCGAGGGAACTGATCCATTAC  
 TGAGGGAACATAATTCAGTCTCATGAGAGAGAGAACTCACTCACTACTGCAAGAATGACAC  
 CAAGCCATTTCATGAGGGATCTGCCTCCGTAACCCTGACACCTCCTGCTAGGTCCCTCCTC  
 CCAACACGGCCACATCAGGGATCAGACTTCAACATGAGTTTTTGTGGGGACAAACAAAAC  
 [G, A]  
 TAGCACTTGCCTTTGCCTTTTGGTTCTATTACATCCTCCACAGGATTGCATTATGCCTAC  
 CCATTTGGTGAGGGCAGTCTCTTTAATTGGTTTACTGATTCAAATGCTACCTCCTCCA  
 GAGACATCCTCACAGACACACCCAGAAATCATGTTTTACCAGTTATCTGGGCATCCCTTA  
 GTCCAGACGAGTTGATACATAAAATTAACCATCACACATGGGATAGAATTAGGATTACAC  
 AGTCAACCTTTATGGGAGAAAATTCAGAGGCATGTCAGGGGTTTATGTAATGTCAAGGA

6455 T TGTATTGTCATTGAGTGGAAATCAGGATTTCACTCCATTAAGTAATTCCTCTGTTAACAA  
 AGAGGGTTCATTTTCATTTTTATTTTCATTAATATTGCTTTTTTTTTTTTTTTCTGGAGAC  
 AGAATCTTGCTCTATCACCAAGGCTGGAGTGCAGTGGTGCATCTCGGCTCACTGCAGCC  
 TCTGCTTCCCTGGATTCAAGCGATTCTTGTGCCTCAGCCTCCCAAGCAGCTGAGATTACAG  
 GCACATGCCACCACACCTGGTTAACTTTTGTATTTCTAGTAGAGATGGGATTTTGCCAT  
 [T, G]  
 TTGGTCAGGCTGGTCTTGAATTCCTGGCCTCTAGTGATCTGCCTGCCTCTGCCTCTGAAA  
 GTGCTAAGATTACAGGCATGAGCTACCATGGCCAGCCCATTTCCTTAATATTTAATTGT  
 CAGACATGTTATGGTTTCTGGCACAATATTAAGAAGACATGATATGAAATCACAGGGTGA  
 ATTTTAGGGCATCACAACAGAAAGATTATGGTATAAGAAAAACAATGGAATTCCAACTAC  
 ATTTCTGTCAAATGTTCTAAAAATATATAAAATCTGTATCTTTTGTGTTCTCTCCTGATTT

6533 T TATTTTCATTAATATTGCTTTTTTTTTTTTTTTTTCTGGAGACAGAATCTTGCTCTATCAC  
 CAAGGCTGGAGTGCAGTGGTGCATCTCGGCTCACTGCAGCCTCTGCTTCCTGGATTCAA  
 GCGATTCTTGTGCCTCAGCCTCCCAAGCAGCTGAGATTACAGGCACATGCCACCACACCT  
 GGTAACTTTTGTATTTCTAGTAGAGATGGGATTTGCCATGTTGGTCAGGCTGGTCTT  
 GAATTCCTGGCCTCTAGTGATCTGCCTGCCTCTGCCTCTGAAAGTGCTAAGATTACAGGC  
 [T, G, A]  
 TGAGCTACCATGGCCAGCCCATTTCCTTAATATTTAATTGTGACACATGTTATGGTTTC  
 TGGCACAATATTAAGAAGACATGATATGAAATCACAGGGTGAATTTTAGGGCATCACAAC  
 AGAAAGATTATGGTATAAGAAAAACAATGGAATTCCAACTACATTTCTGTCAAATGTTCT  
 AAAATATATAAAATCTGTATCTTTTGTGTTCTCTCCTGATTTATATTCTAAATTTGATGT  
 TATCCTTCTCTGCAGAAATAAGTGTCTGAAAGATGAAAAAATGGAAGAATTCTTTAG

6919 A TGAAATCACAGGGTGAATTTTAGGGCATCACAACAGAAAGATTATGGTATAAGAAAAAC  
 AATGGAATTCCAACTACATTTCTGTCAAATGTTCTAAATATATAAAATCTGTATCTTTT  
 GTGTTCTCTCTGATTTATATTCTAAATTTGATGTTATCCTTCTCTGCAGAAATAAAGTG  
 TCTGAAAGAATGAAAAAATGGAAGAATCTTTAGTAAGGTATAAAATACCTTTCTATC  
 TTTGTAGCATTCTAAGCCTTTTGTACCTTTCCAACTCCCAACATGCCATATTCCTGA  
 [G, C]  
 TAGGCCACAGCCATGTACATTGATCCCTTTATTTTCTTCTCTCTGCCTGAGATTTCTCTC  
 ATTCCCCCTTCTCTGCCTGGTATATGATTGCCCATGTTTAAAGCCCCAACTCACCTTTA  
 TAATCTTCTAGCCACTTCTTTATCGGTATTCCAGAAAAACAAAAGAAGCTTCCACA  
 AGACAACATCTGTAATACACTGCTTAACTTCTTTTGACCCTGCTGAGTTCAAAAATCTT  
 ATCTTTTAAAGGATTGAATGGAGTCCACCAAGGTATCTATATTGACAGGATTTATGAAA

7305 G ATTGCCCATGTTTTAAGGCCCCAACTCACCTTTATAATCTTCTAGCCCACTTTCTTTA  
 TCGGTATTCAGAAAAACAAAAGAAGCTTCCACAAGACAACATTCTGTAATACACTGCT  
 TAACTTCTTTTGACCCTGCTGAGTTCAAAAATCTTATCTTTTAAAGGATTGAATGGAGTC  
 CACCAAGGTATCTATATTTGACAGGATTTATGAAAACAAAAGGATTTGTTGAGAAAGTTT  
 GAAGCCTAACTCTGAAACGTGGATCATAGTGTCTTACTACACATTAAGTGTCTTAGTGGAT

FIGURE 3, page 27 of 42

[G, A]  
TAATAGTTATTATTATAGGCTGTGGAATCAGAACAGGGTTCAAATGTTTTACCGCTTGC  
TAGACTGTGGCCTTGGGCATGTTATTTAATGCCTGGAGGCCTCAAATGTTAACTAGGAAT  
GGTAAGACCTACCCAGTAACTTAGCATAAATAGTAAATTCATTCTAATGTTTTCAA  
CAGTGCCAGACATTGTTTAACTGAGGATATAGTGGTGAACAACACTGACAGCGTTC  
TTCATTGTATTCTCAAAACCCCTCCCTATAGTAAGTAGGTCTGTGTGTGTGTAGGTGCA

7340 TAATCTTCCTAGCCCACTTTCTTTATCGGTATTCCAGAAAAACAAAAGAAGCTTCACA  
AGACAACATTTCTGTAATACACTGCTTAACTTCTTTGACCCTGCTGAGTTCAAAAATCTT  
ATCTTTTTAAGGATTGAATGGAGTCCACCAAGGTATCTATATTTGACAGGATTTATGAAA  
ACAAAAGGATTTGTTGAGAAAGTTGAAGCCTAACTCTGAAACGTGGATCATAGTGTTTA  
CTACACATTAACCTGTTTTAGTGGATGTAATAGTTATTATTATAGGCTGTGGAATCAGAAC  
[A, G]  
GGGTTCAAATGTTTTACCGCTTGCTAGACTGTGGCCTTGGGCATGTTATTTAATGCCTG  
GAGGCCTCAAATGTTAACTAGGAATGGTAAGACCTACCCAGTAACTTAGCATAAATAGTA  
AATTCATTCTAATGTTTTCAAACAGTGCCAGACATTGTTTAACTGAGGATATA  
GTGGTGAACAACACTGACAGCGTTCTTCATTGTATTCTCAAAACCCCTCCCTATAGTAAGT  
AGGTCTGTGTGTGTGTAGGTGCATGGGGAATAAAAAATAAAGCAAATAATGAACAG

7466 TTAAGGATTGAATGGAGTCCACCAAGGTATCTATATTTGACAGGATTTATGAAAACAAAA  
GGATTTGTTGAGAAAGTTGAAGCCTAACTCTGAAACGTGGATCATAGTGTTTACTACAC  
ATTAAGTGTTTAGTGGATGTAATAGTTATTATTATAGGCTGTGGAATCAGAACAGGGTT  
CAAATGTTTTACCGCTTGCTAGACTGTGGCCTTGGGCATGTTATTTAATGCCTGGAGGC  
CTCAAATGTTAACTAGGAATGGTAAGACCTACCCAGTAACTTAGCATAAATAGTAAATTC  
[A, G]  
TTCATTTAATGTTTTCAAACAGTGCCAGACATTGTTTAACTGAGGATATAGTGGTG  
AACAACACTGACAGCGTTCTTCATTGTATTCTCAAAACCCCTCCCTATAGTAAGTAGGTCT  
GTGTGTGTGTGTAGGTGCATGGGGAATAAAAAATAAAGCAAATAATGAACAGGGTAAT  
TTCAAAAAGCAGAAAGAGCTATTCAACAAAACCTACCTGCCTTTTATTAGATGAAACTCTC  
AACTCTATGGTTTGTCTCTCCTGTCAATTCTGTTAAATGCTGTCAGCCTGTTTTCTCTTA

7589 AACTGTTTTAGTGGATGTAATAGTTATTATTATAGGCTGTGGAATCAGAACAGGGTTCAA  
ATGTTTTACCGCTTGCTAGACTGTGGCCTTGGGCATGTTATTTAATGCCTGGAGGCCTC  
AAATGTTAACTAGGAATGGTAAGACCTACCCAGTAACTTAGCATAAATAGTAAATTCATT  
CATTTAATGTTTTCAAACAGTGCCAGACATTGTTTAACTGAGGATATAGTGGTGA  
CAACACTGACAGCGTTCTTCATTGTATTCTCAAAACCCCTCCCTATAGTAAGTAGGTCTGT  
[G, C]  
TGTGTGTGTAGGTGCATGGGGAATAAAAAATAAAGCAAATAATGAACAGGGTAATTTT  
AAAAAGCAGAAAGAGCTATTCAACAAAACCTACCTGCCTTTTATTAGATGAAACTCTCAAC  
TCTATGGTTTGTCTCTCCTGTCAATTCTGTTAAATGCTGTCAGCCTGTTTTCTCTATCA  
CCCTGGCCACGACTTCTGTCTTTTCTGCTTGGTCTGTAGACTCTAACCCAAGGCTCATT  
CTCTGCCTGGCTATCTGCCTTCTGTGGCTCTTTGCCACTACCTACATTTTCTGTGTGCA

7810 CTGGGGATATAGTGGTGAACAACACTGACAGCGTTCTTCATTGTATTCTCAAAACCCCTCC  
CTATAGTAAGTAGGTCTGTGTGTGTGTAGGTGCATGGGGAATAAAAAATAAAGCAA  
ATAATGAACAGGGTAATTTCAAAAAGCAGAAAGAGCTATTCAACAAAACCTACCTGCCTTT  
TATTAGATGAAACTCTCAACTCTATGGTTTGTCTCTCCTGTCAATTCTGTTAAATGCTG  
TCAGCCTGTTTTCTTATCACCTGGCCACGACTTCTGTCTTTTCTGCTTGGTCTCTGTAG  
[A, C]  
CTCTAACCCAAGGCTCATTCTCTGCCTGGCTATCTGCCTTCTGTGGCTCTTTGCCACTAC  
CTACATTTTCTGTGTGTCACAGGAAGGACCATTCCTGTGGACCATAAAATCTCTTTT  
TGAAAGAATTCTTCTGATTGGGCCACAGCACATCTTGGAACAGCATTAGACATTTG  
CCACTGCTCAGCAGCTCTGGGGGAAAATGTTTACTGAGAAGCGTACAGTAGTTTTTTTGA  
CTAACCATGGTGCAACCTCCTCCAGAGGGAAACCTATGAGTATTTCAAGGACATGTGAT

9104 TTAAACGAATTATTGTAGAAACAGAAAAACAAATACTGTGTTCTCATTACAGGGGGAGC  
TAAACCTTGGGTAAATGGGGCATAAAGATGGGAACAATAGACACTAGGGACTCCAAAAGG  
GGGGAGGGAGGGAGGGGCAAGGGCTGGAAAGCTTCTACTGGGTACTTTGTTCAACAC  
CTGGGTGATGGCAGATTAGGAGCTCAAACCCAGTATCACACAGTATACCCTTGTAACA  
AGCTGATGGTGTAACCCCTGAATCTACAATAAAATTATTTTATTTTAAAAAATCATTATA  
[G, A]  
GGATTTTTTAAAAAGAAGGATTCTTAGACAGGTGCAGCCAAACAATTTTTTTTAAATGTTG  
GCAGGCCGCCACCGCCAGTCACTTATGCTGCAATAGCCCATGTCCCAACATTCCCAACCT

FIGURE 3, page 28 of 42

ACTTCTCTCCAAAAGAGAAGCTATACTTTTCAGATGGCCCTGTGCTGGGTTCTCCCTGGAA  
GTTTCTGGGAAAGGGGCTTGAGTTGCCCGACTGGACTCTTCTGGAGTGGGAGCCGGG  
GCTTCTGATCAGACGTGAGTGAGGCAGGAACCCGCGGTCTCCAGCGCAGCCAGAGTG

9503 CATGTCCCAACATTCCCAACCTACTTCTCTCCAAAAGAGAAGCTATACTTTTCAGATGGCC  
CTGTGCTGGGTTCTCCCTGGAAGTTTCTGGGAAAGGGGCTTGAGTTGCCCGACTGGAC  
TCTTCTGGAGTGGGAGCCGGGCTTCTGATCAGACGTGAGTGAGGCAGGAACCCGCGG  
TCTCCAGCGCAGCCAGAGTGCGGTCCACGCAGTCCCGGGTCTGCGCGCTCGCGCC  
TTTGCGCTGAAGCCGTTAGGATGAGCCCTCTCTTCCAGAGCTTTAACCGATGAAGGTGC  
[A, T]  
TTGTGTTTGGCGCCCTGAGGAGGATGCTGTCTTAGGCCTCTTCCCACTGGACGTGTGTG  
GTGGGCAGAGATCCCGTTCGTGCGTCCGACTTCCACCCGCTGGGGCTCACTCAGGCCGC  
GGAGCTGCGAGGGAGACATCCTCGATGGACTCCCTCTACGGAGATCTCTTTTGGTACCTG  
GACTATAACAAGGATGGGACCTTGGACATTTTTGAGCTTCAGGAAGGCCTGGAGGATGTA  
GGGGCCATTCAATCTCTAGAGGAAGCGAAGGTGGGTCTCACTGGGGCTGTAATCAGAGAG

9898 ACCCCGCTGGGGCTCACTCAGGCCGCGGAGCTGCGAGGGAGACATCCTCGATGGACTCCC  
TCTACGGAGATCTCTTTTGGTACCTGGACTATAACAAGGATGGGACCTTGGACATTTTTG  
AGCTTCAGGAAGGCCTGGAGGATGTAGGGGCCATTCAATCTCTAGAGGAAGCGAAGGTGG  
GTCTCACTGGGGCTGTAATCAGAGAGACGTTGGGGCTGGGAGCCCTGGAGAGGCATTGGG  
CAGAGAGGGCAAAATTTACATGTTGTCAAGCTTGACCTGGGCCCCTGCAGTGTTCAAGT  
[G, C]  
GTTGACCAGCGTTACCGTTTATTAAGAATAACAACACAGCTAACACATTTCTCAAGTATT  
TTTCTCCGTTTTCTCCTTGGCTGTAGTAAATCTCCAACCTCAGATTGCTCTCAAGATGT  
TGGCTACATACAGCCTTGTCTTAGGAGTCACCTTGTTCAATGTGCTCACCTGTCATTAGT  
CACCCAGAGGGCGCTCTAGGCTAAAGATGCGCCCTCCCCAGTTCAGAGAACTGGAATAAT  
CACTCTACGTGTATTTGGGAGTGGGGTGGTGATTGGAATTTTCTGATGTTATGTTTGG

10196 GTGGTTGACCAGCGTTACCGTTTATTAAGAATAACAACACAGCTAACACATTTCTCAAGT  
ATTTTTCTCCGTTTTCTCCTTGGCTGTAGTAAATCTCCAACCTCAGATTGCTCTCAAGA  
TGTTGGCTACATACAGCCTTGTCTTAGGAGTCACCTTGTTCAATGTGCTCACCTGTCAAT  
AGTCACCCAGAGGGCGCTCTAGGCTAAAGATGCGCCCTCCCCAGTTCAGAGAACTGGAAT  
AATCACTCTACGTGTATTTGGGAGTGGGGTGGTGATTGGAATTTTCTGATGTTATGTTT  
[T, C]  
GGTTTCTGTTTCTGGAAGGGGCGAGTGAAGTGGCTTTTACTCTCGGGTTTCACTAGTGC  
TGAGGTTTCTCATAATATGCCCTTAATTGATAGACCCTAGTTATCAGTACCGAGCTTAGG  
CTAACCCCTTCTCTTCCCCAGAAGGCTAACCTACAGGCTCCTTCTCAGCATGTTGTGCTTC  
GTACATACTCCTATTGCAGTATTTCCAAGTCATTTTTCAATTGGAATTTATTATTGTATA  
TAATAATTACTTTATAAGTATATTTGCTCTTTGGATGTTTGACCCGGTAGACTGGGAGAT

12327 GTCATGTTATTTAATGCCTGGAGGCCTCAAATGTTAACTAGGTAATGGTAAGACCTACCC  
AGTAACTTAGCATAAATAGTAAATTCATTCAATTAATGTTTCAAACAGTGCCAGACATT  
GTTTAAATGAAGTGGGATATAGTGGTGAACAACACTGACAGCGTTCTTCATTGTATTCTC  
AAAACCCCTCCCTATAGTAAGTAGGTCTGTGTGTGTGTGTAGGTGCATGGGAATAAAAAA  
TAATAAGCAAATAATGAACAATAAAATTATTTATTTAAAAAAAAGAAATGATACTTAC  
[C, G, A]  
TTGTCGTGTTAAGATACAAAAGCAATAACTTTTTATTGTGAAAATAGTCTGTTTTTGAAC  
AATATATTGTTTTGTTTTTCTGTGAAAGTTGAGAACTAAATATACGAAGAGATAATG  
GTCAGACCATAAATAAAATAGAACTTTGACTCAAAATTTACAGCAGTCTGCCAGAAAA  
CCAGCCCTTTATCTAAATAAACAGACCAGGAACACAGCCTGTTATGTGAGACTTATAGG  
AAGTCAGGTTGCTATCTCTAGAGACAATACACAAAGCTATGCAATAACTGCTGTAAACAGC

13749 TACAGGCGTGAGCCACCATGCGCCAGCCATAGACTATATATTTTTGATCTGATAACTGG  
TTCAGCTACTAAGTGACTAACAGGCAAGTAGCATCTATAGTGTGGATATGCTGGACAAAA  
GGACATTCACTCCTGGGCAGGATGGCACAGAATGTTGAGAGATTTTATCATGCTACTCA  
GAATGGTGTGCAATTTAAACTTATGAGTTGTTTGTCTGGAGTTTTCCATTTAATAGT  
TCAGACCATGGATTGACCGCAGGTAACGTGAACTGTGGAGAGTGAACTGTGGATAAGGG  
[G, A]  
GGACTATTGTATTGTTAAGTCAGACTCATTAGGCAATCATAACTCTTGATTTGCCATCAG  
AAATGCTGCAGAAATATGGGTAAAAAACTGTTCAAAAATAGGGTCAGGGATGTCCTT  
TAACTTGTTACTTCCAAAATGTTAGTGAAGTGTGGCCCCAAAGAGTGAAAGGAACAAA  
TGACTAAGAGAAAATCTGTTTTTCAAGATGACAGATTAAAAAGAAAGCAACTTGCTGAAA  
CACTGAAAATCTCTCACTTGTAAGATAACACAAAACCTGGCTAAAACCTGGTTGGAATGAA

FIGURE 3, page 29 of 42

14150 ATAGGGTCAGGGATGTCCTTTAACTTGTTACTTCCAAAATGTTAGTGAAAACCTGTGGCCC  
CAAAGAGTGAAAGGAACAAATGACTAAGAGAAAATCTTGTTCAGGATGACAGATTAAA  
AAAGAAGCAACTTGCTGAAACACTGAAAATCTCTCCACTTGTAAGATAACACAAAACCTGG  
CTAAAACCTGGTTGGAAATGAATATGGCCAACTCAAGTCTGCACAGAACTAAGTTGGTGATG  
TTACAGCCCCAAATTTCCACCACATATTTTATACTAACTCCCCCGGATTTTCACACATGA  
[T, C]  
CTGTGAGGTAGCATGAAGAGGTAACATATGCATGCCTAAGGACTTGGGAGACCTCCCCATT  
TCCTTCCACCAATCACCCTAATCCCGAATCCGCCCCAAACCTTTTCTAATAACTAC  
CTTAAAGCCAGCATAGGGAGACAGATTTGAGCTGGACTCCTGTCTTCTTGTGGGTACCT  
TGCAATAAAAAGCTTTTCTTTCTCAACACCTGGTATTATAGTATTGACTTCTAGTTTCTAT  
CGGGCAGCAAGCCCCCTTTGGTGGTGACTATTCTTGTTCGCTGATATTTCCATTGGCCA  
14529 ACTAATCCCAGAATCCGCCCCAAACCTTTTCTAATAACTACCTTAAAGCCAGCATAGGG  
AGACAGATTTGAGCTGGACTCCTGTCTTCTTGTGGGTACCTTGCAATAAAAAGCTTTTC  
TTTTCTCAACACCTGGTATTATAGTATTGACTTCTAGTTTCTATCGGGCAGCAAGCCCCCTT  
TGGTGGTGACTATTCTTGTTCGCTGATATTTCCATTGGCCAAAATATAAACCTCTTAGA  
TGAAACTTCAGTACGTAAATGGCGCCACAGAATGCTGTGACATTTTCTCTTGATTATA  
[G, A]  
CAGGTTACTTTACTGAATACCGTAGGCAGTTATAACACACTAAGTATTTGTGTATCTAAA  
CATAGAAAAGATACAGTAAAAATATGGTAATTTTTTTCAACTTTTAGTTGAGATTGGAG  
GGTATGTGCACATTTGTTACAAGGGTATATTGCATGATGCTGAGGTTTGGGGTACAATTG  
AACCCTGTCAACCAGGTAGTGAGCATAGTACCCAATCGATAATTTTCAACCCTGTCCA  
TTCCCTCCCCGTTCTTGTAGTCCCCAGTTTCTGCTTTTCCATCTTTATATCCGTGTGCA  
14653 CTCAACACCTGGTATTATAGTATTGACTTCTAGTTTCTATCGGGCAGCAAGCCCCCTTTGGT  
CGGTGACTATTCTTGTTCGCTGATATTTCCATTGGCCAAAATATAAACCTCTTAGATGAA  
ACTTCAGTACGTAAATGGCGCCACAGAATGCTGTGACATTTTCTCTTGGATTATAGCAG  
GTTACTTTACTGAATACCGTAGGCAGTTATAACACACTAAGTATTTGTGTATCTAAACAT  
AGAAAAGATACAGTAAAAATATGGTAATTTTTTTCAACTTTTAGTTGAGATTGGAGGGT  
[G, A]  
TGTGCACATTTGTTACAAGGGTATATTGCATGATGCTGAGGTTTGGGGTACAATTGAACC  
CTGTCAACCAGTAGTGAGCATAGTACCCAATCGATAATTTTCAACCCTGTCCATTCC  
CTCCCCGTTCTTGTAGTCCCCAGTTTCTGCTTTTCCATCTTTATATCCGTGTGCACCCC  
ATGTTTTGCTCCCATGTGTATGTGAGAACTTGTGGTGTGGTTTTCTATTTCTGCGTTG  
ATTCGCTTAGGATAATGGCCTTCAGCTGCATCCATGTTGCTGCAGAGGACGTGATTTTAT  
15871 AGGAGTTTATCAATTTTATTAGTCTTTTCAAAGAACCATCTTTGGCTTTGTTAATCCTC  
CCAATGGTGTGTTTTCTTTCTCATTACTTTTTGCTCTTTATTTCTTCAACTTCTTTTTT  
GCTTAATTTTAAATAATTTCTTGAGATTGAGATAAGCCTCAATGATGGGTACCGGATTT  
CCAGTCTTTCTTTCTTTCTAATTATGCATTTTAAACCAGAAATCTTTCTTAAGTGTAGC  
TTTAGTTGCAGCTCACAAGTTTTCAGATCTGTCTCTCAGTCTGGAGGTTGGAGATCTGACC  
[A, G]  
TGACCATGAAACCATCCAGTCACAATGTGGCATTATTTTTTTAATTTTTTTTTTTTTTT  
TGAGATAGAGTTTCACTCTTATTGCCTAGGCTGGTGTGCAATGGTGCGATCTCGGCTCAC  
AGCAACCTCCACCTCCCAGGTTCAAGCGATTCTTTGCCTCAGCCTCCCAAGTAGCTGGG  
ATTACAGGCATGCGCCACCATGCCCCAACTAATTTGTATTTTATAGTAGAGATGGGGGTTT  
TCCATGTTGGTCAGGTTGGTCTTGAACCTCCGACCTCAGGTGATCCGCCACCTCAGCCT  
19244 GTGGCATTATTGGTTCATATTTTTATTTTTTAGACTTCCTTAATGCAAAACATATACAGT  
TGATCCTCATTATTTGGGGATTCTGTATTTGCAAATTTGCCTACTCAATAAAATTTATCC  
CCAAAGTAACCCCAAAATATATACTCACAGTACTTTCCAGGCATTATGGACATGCACA  
GAGCAGTGAAAACTTGAGTTGCTCAGCATGTACATTCTAGCTAGTAGAATAAGGCAAT  
ACTCTGCCTTCTTGTTCAGCTCTCATACTATTAAGTAGCAAGTATCCCTTTCAAGGTCT  
[G, A]  
TTTTGTGCCAGTTTTTGCATTTTTGTATTTTTGTGGTAATTTCTTTTTTAAATGTTCC  
CCAAAGGTAGTGCTGAAGTGCTGTAGTGTTCCTAAGTGCAAGAAAGCCATAGCATGCC  
TTATGGAGAAAATATATGCGTTGGATAAGCTTTGCCCAAAATCAATGTTAGTGAATCAA  
CAGCACACATTAAATGAGGTGCCTTCAAACAGAAACAGACATAAGACATGGTTATGTATT  
AATCAGTTGATGAAAGTGTGTAATCAGAGGCTCACAGGAACCTAACCTGTTTTCTCTG  
19387 CTCACAGTACTTTCCAGGCATTTCATGGACATGCACAGAGCAGTGAAAACTTGAGTTGC  
TCAGCATGTACATTCTAGCTAGTAGAATAAGGCAATACTCTGCCTTCTTGTTCAGCTC

FIGURE 3, page 30 of 42

TCATACTATTAAGTAGCAAGTATCCCTTTCAAGGCTCTATTTTGTGCCAGTTTTTGCATTT  
TTGTATTTTTTGTGGTAATTTCCCTTTTAAAATGTTCCCAAAGGTAGTGCTGAAGTGCT  
GTCTAGTGTTCCCTAAGTGCAAGAAAGCCATAGCATGCCTTATGGAGAAAATATATGCGTT  
[T,G]  
GATAAGCTTTGCCCAAATTC AATGTTAGTGAATCAACAGCACACATTAAATGAGGTGCC  
TTCAAACAGAAACAGACATAAGACATGGTTATGTTATTAATCAGTTGATGAAAGTGTTGTA  
ATCAGAGGCTCACAGGAACCTAACCCTGTTTTCTCTAGGAACAATGGTTTGGTATTGG  
CTAATTGCTGTTTTGCAATGAATATAGAACTTTATGGAAGATGATTGCTGTGAATAATGA  
GAATTAACCATATCTCTTTAAGAGTGCATTTCTAAAGGAGAATATTCAGAAGGGTATTGG

19447 TCAGCATGTACATTCTAGCTAGTAGAATAAGGCAATACTCTGCCTTCTTGTTTCAGCTC  
TCATACTATTTAACTAGCAAGTATCCCTTTCAAGGTCTATTTTGTGCCAGTTTTTGCATTT  
TTGTATTTTTTGTGGTAATTTCTTTTTTAAAATGTTCCCCAAAGGTAGTGCTGAAGTGCT  
GTCTAGTGTTCTAAGTGCAAGAAAGCCATAGCATGCCTTATGGAGAAAATATATGCGTT  
GGATAAGCTTTGCCCAAATTCATGTAGTGAATCAACAGCACACATTAATGAGGTGC  
[C, G]  
TTCAAACAGAAACAGACATAAGACATGGTTATGTATTAATCAGTTGATGAAAGTGTTGTA  
ATCAGAGGCTCACAGGAACCTAACCCCTGTTTTTCTGTAGGAACAATGTTTTGGTATTTG  
CTAATTCAGTGTTTGCAATGAATATAGAACTTTATGGAAGATGATTGCTGTGAATAATGA  
GAATTAACCATATCTCTTTAAGAGTGCATTTCTAAAGGAGAATATTGAGAAGGGTATTTG  
CATAATTTCTTTACTAACAGATGCTGCCTCTCACTGTCCTTACATGGTCCAGATTCTCAT

[illegible]

20492 CACTCTGTCAACCCAGGCTGGGGTGCAGTGGCGGATCTCAGCTCACTGCAACCTCCGCCT  
CCCAGCTTCAAGCAATTTTCTCTTTTAGCCTCCCAAGTAGCTGGGATTATGGGCACCTG  
CCACCATGCCTGGCTGATTTTCTATTTTCTAGTAGAGACGGGGTTTACCATTGTTGGCCA  
CACTGCTCTCTAACTGCTGACCTCAGGTGAACACCCGCCTCAGCCTCCAAAAGTGCTGG  
GATTGCAGGTGTGAGTCACTGTGCCTGGCCTTTTACTGCCAGATTTTTAAAGAATAGTC  
[T, -]  
GTGCTTTAGCTCTATTTTCTCATTACTACTTCTCTTTAACTCAGTCATATATGATGTTT  
TGCATAGTAAATGTCTAGTAATTTATTAATAATGTAGAAATAGGTACTTTTTAAATGAAT  
AGATCCTACTTTTAAATGAATTTATCTTGGAGTTAGAATATCTTGATTTGGATTTTAGTTC  
TGCTACTTCTTAATTACATTACTTGGTAAGGCCACTTGTGAAGTCAGTCTCTTTGGAGGA  
ATATTATTTATCTATAAGGCTGTTACAATTACTGAATTTTAAAAAATGTGTATTTATTTT

20868 TAGTAATTTATTA AAAATGTAGAAATAGGTACTTTTTAAATGAATAGATCCTACTTTAAT  
TGAATTTATCTTGGAGTTAGAATATCTTGATTGGATTTTAGTTCTGCTACTTCTTAATT  
ACATTACTTGGTAAGGCCACTTGTGAAGTCAGTCTCTTTGGAGGAATATTATTTATCTAT  
AAGGCTGTTACAATTACTGAATTTTAAAAATGTGTATTTATTTTAAATGTATTTGTTA  
CATTTTTAGTATTGATGTTGGGATAGGCATTTAAGCAAGTCTATAACTCACCTACATGCA  
[T,C]  
AATTTTGCCTTAATCAGTTTAAAGCTTTCTCTTAAATGAGAGATTTGAAATTCATAATTT  
CTGTGGTTCCTATCAGTTCTGAGTTTATTTTTTGCCCTTTTTATTTTTTAAAGGAAAA  
ATTGAGGCTTCAGAAATGTCCAGTCTCTCCAGACACTGGGTCTGACTATTTCTGAACAA  
CAAGCAGAGTTGATTCTTCAAAGGTAAGCTCTTCATGTTGGTCAACAATTGACTTTCACT  
TTAATATCCTGCATTAGAACTCTGTGTTTGTAAGTGTGGCTTTAAACACCTCCCTAGTC

20941 GAGTTAGAATATCTTGATTGGATTTAGTTCTGCTACTTCTTAATTACATTACTTGGTA  
AGGCCACTTGTGAAGTCAGTCTCTTTGGAGGAATATTATTTATCTATAAGGCTGTTACAA  
TTACTGAATTTTAAAAAATGTGATTTATTTTAAATGTATTGTTACATTTTGTGATT  
GATGTTGGGATAGGCATTTAAGCAAGTCATAACTCACCTACATGCATAATTTTGCCTTA  
ATCAGTTTAAAGCTTTCTCTCTAAATGAGAGATTTGAAATTCATAATTTCTGTGGTCTTA

[T,C]  
CAGTTCTGAGTTTTATTTTTTGGCCTTTTATTTTTTAAAGGAAAAATTGAGGCTTCAG  
AAATTGTCCAGTCTCTCCAGACACTGGGTCTGACTATTTCTGAACAACAAGCAGAGTTGA  
TTCTTCAAAGGTAAGCTCTTCATGTTGGTCAACAATTGACTTTTCACTTTAATATCCTGCA  
TTAGAACTCTGTGTTGTAAAGTGTGGCTTTAAACACCTCCCTAGTCTTCATTATGTATA  
TCCAAGATCTTTTTGTCTTTTTTCCCTCCCATTCATTTTGTATGTGTACATTTATCTAAAG

21116 GTATTGATGTTGGGATAGGCATTTAAGCAAGTCTATAACTCACCTACATGCATAATTTTG  
CCTTAATCAGTTTTAAAGCTTTCTCTTAAATGAGAGATTGAAATTCATAATTTCTGTGGT  
TCTTATCAGTTCTGAGTTTTATTTTTTGGCCTTTTATTTTTTAAAGGAAAAATTGAGG  
CTTCAGAAATTGTCCAGTCTCTCCAGACACTGGGTCTGACTATTTCTGAACAACAAGCAG  
AGTTGATTCTTCAAAGGTAAGCTCTTCATGTTGGTCAACAATTGACTTTCACTTTAATAT  
[C,T]  
CTGCATTAGAACTCTGTGTTGTAAAGTGTGGCTTTAAACACCTCCCTAGTCTTCATTAT  
GTATATCCAAGATCTTTTTGTCTTTTTTCCCTCCCATTCATTTTGTATGTGTACATTTATC  
TAAAGTGTAAAGATGGGAAGTGTAAAGCTCAGACTGGACTCTTTCTTTCAAGGCCTCAAAG  
GATAGTGAAGTGGCAGGAAGTAAAGGTTTTAACTCCATAGATGAGGAGCTGAAGAGTTTTG  
GTGTTGCTTTTTCTCCATTTGATTTCTAATGTGACAGTAAACTCATTGATTCAAACTAA

21701 CATTGATTCAAACTAAGAAGACTAGCAGATTCATCACATTATTTAACCTAGATGTGACTG  
GAAAAAAGGGAAATTACTAAGCTCTCCAAGCTAACAAAGAAATACCTGTTTAACTTTTCA  
GAAACAGAAATGCAATTTGAACCTTATTGTCTGGGGCAATCAGTTTGACTATTTAAGT  
CAGACTTTTATACTCTTAATGTTTTGTTTCATGGGATAGAGCAGTAATCTCTGCAGCCCA  
GGTGCTCTCAAATACTCTGTTGCTATAAACACAGGGCAGGAAGTATTTTTATGATAAC  
[G,A]  
TAAACAGAAAAAGGACAATTATATTGTATTAATATTGTTGTGAATATTTTCAGTCCTCAC  
ATTGTCTAAAAATCTTTCTAAATGGCTTTGTTATTGAATTTATCTCATTTTATATCTGTG  
CCAACAGCATTTTCATCCTTTCTCTTATAATTTCTTTTACAAACAGCTGCTCAAGAGGA  
AGGCTCAAAGTCTCAAGGCTGAGCACGTAATGACTTTTGTAGTACTAGATGAGAAGGGC  
TTTCTGAGGAAATGAAAACCTAAAACATGAAAAGAAGATAAACAGAAATTTGGACAGTGA

21710 AAATAAGAAAGACTAGCAGATTCATCACATTATTTAACCTAGATGTGACTGGAAAAAAGG  
GAAATTACTAAGCTCTCCAAGCTAACAAAGAAATACCTGTTTAACTTTTCAAAAAACAGA  
AATGCAATTTGAACCTTATTGTCTGGGGCAATCAGTTTGACTATTTAAGTCAGACTTTT  
ATACTCTTAATGTTTTGTTTCATGGGATAGAGCAGTAATCTCTGCAGCCAGGTGCTCTC  
AAATACTCTGTTGCTATAAACACAGGGCAGGAAGTATTTTTATGATAACGTAAACAG  
[A,-]  
AAAGGACAATTATATTGTATTAATATTGTTGTGAATATTTTCAGTCCTCACATTGTCTAA  
AAATCTTTCTAAATGGCTTTGTTATTGAATTTATCTCATTTTATATCTGTGCCAACAGCA  
TTTTTCATCCTTTCTCTTATAATTTCTTTTACAAACAGCTGCTCAAGAGGAAGGCTCAAA  
GTCTCAAGGCTGAGCACGTAATGACTTTTGTAGTACTAGATGAGAAGGGCTTTCTGAG  
GAAATGAAAACCTAAAACATGAAAAGAAGATAAACAGAAATTTGGACAGTGAATATAGAG

21826 CAGAAATGCAATTTGAACCTTATTGTCTGGGGCAATCAGTTTGACTATTTAAGTCAGAC  
TTTTATACTCTTAATGTTTTGTTTCATGGGATAGAGCAGTAATCTCTGCAGCCAGGTGC  
TCTCAAATACTCTGTTGCTATAAACACAGGGCAGGAAGTATTTTTATGATAACGTAAA  
ACAGAAAAGGACAATTATATTGTATTAATATTGTTGTGAATATTTTCAGTCCTCACATTG  
TCTAAAATCTTTCTAAATGGCTTTGTTATTGAATTTATCTCATTTTATATCTGTGCCAA  
[C,T]  
AGCATTTTCATCCTTTCTCTTATAATTTCTTTTACAAACAGCTGCTCAAGAGGAAGGCT  
CAAAGTCTCAAGGCTGAGCACGTAATGACTTTTGTAGTACTAGATGAGAAGGGCTTTCC  
TGAGGAAATGAAAACCTAAAACATGAAAAGAAGATAAACAGAAATTTGGACAGTGAATAT  
AGAGCATATAATATTCTGCTTCTAAAGTAATATTCTTCTAGGAAAGTGAGGGCGTTCCC  
TGGCTGTTAGGCCAGAAATCATATTTCTATATTTCTTTGATAGCTTTAGGAATAATGCA

21840 TGAACCTTATTGTCTGGGGCAATCAGTTTGACTATTTAAGTCAGACTTTTATACTCTTAA  
TGTTTTGTTTCATGGGATAGAGCAGTAATCTCTGCAGCCAGGTGCTCTCAAATACTCTG  
TTGCTATAAACACAGGGCAGGAAGTATTTTTATGATAACGTAAAACAGAAAAGGACAA  
TTATATTGTATTAATATTGTTGTGAATATTTTCAGTCCTCACATTGTCTAAAAATCTTTC  
TAAATGGCTTTGTTATTGAATTTATCTCATTTTATATCTGTGCCAACAGCATTTTCATCC  
[-,T]  
TTCTCTTATAATTTCTTTTACAAACAGCTGCTCAAGAGGAAGGCTCAAAGTCTCAAGGC  
TGAGCACGTAATGACTTTTGTAGTACTAGATGAGAAGGGCTTTCTGAGGAAATGAAAA

CCTAAACATGAAAAGAAGATAAACAGAATTGGACAGTGAGATATAGAGCATATAATAT  
TCTGCTTCTAAAGTAATATTCTTCTAGGAAAGTGAGGGCGTTTCCCTGGCTGTTAGGCCA  
GAAATCATATTCTATATTTCTTTGATAGCTTTAGGAATAATGCAAATCTAAGCCCAA

21841 GAACCTTATTGTCTGGGGCAATCAGTTTGACTATTTAAGTCAGACTTTTATACTCTTAAT  
GTTTTGTTTCATGGGATAGAGCAGTAATCTCTGCAGCCAGGTGCTCTCAAATACTCTGT  
TGCTATAAACACAGGGCAGGAACGATTTTTTATGATAACGTAAAACAGAAAAGGACAAT  
TATATTGTATTAATATTGTTGTGAATATTTTCAGTCCTCACATTGTCTAAAAATCTTTCT  
AAATGGCTTTGTTATTGAATTTATCTCATTTTATATCTGTGCCAACAGCATTTTCATCCT  
[-, C, T]  
TCTCTTCATAATTTCTTTTACAAACAGCTGCTCAAGAGGAAGGCTCAAAGTCTCAAGGCT  
GAGCAGTAAATGACTTTTGTAGTACTAGATGAGAAGGGCTTTCCTGAGGAAATGAAAAC  
CTAAACATGAAAAGAAGATAAACAGAATTGGACAGTGAGATATAGAGCATATAATATT  
CTGCTTCTAAAGTAATATTCTTCTAGGAAAGTGAGGGCGTTTCCCTGGCTGTTAGGCCAG  
AAATCATATTCTATATTTCTTTGATAGCTTTAGGAATAATGCAAATCTAAGCCCAAG

21843 ACCTTATTGTCTGGGGCAATCAGTTTGACTATTTAAGTCAGACTTTTATACTCTTAATGT  
TTTTGTTTCATGGGATAGAGCAGTAATCTCTGCAGCCAGGTGCTCTCAAATACTCTGTG  
CTATAAACACAGGGCAGGAACGATTTTTTATGATAACGTAAAACAGAAAAGGACAATTA  
TATTGTATTAATATTGTTGTGAATATTTTCAGTCCTCACATTGTCTAAAAATCTTTCTAA  
ATGGCTTTGTTATTGAATTTATCTCATTTTATATCTGTGCCAACAGCATTTTCATCCTTT  
[-, C]  
TCTTCATAATTTCTTTTACAAACAGCTGCTCAAGAGGAAGGCTCAAAGTCTCAAGGCTGA  
GCACGTAATGACTTTTGTAGTACTAGATGAGAAGGGCTTTCCTGAGGAAATGAAAACCT  
AAAACATGAAAAGAAGATAAACAGAATTGGACAGTGAGATATAGAGCATATAATATTCT  
GCTTCTAAAGTAATATTCTTCTAGGAAAGTGAGGGCGTTTCCCTGGCTGTTAGGCCAGAA  
ATCATATTCTATATTTCTTTGATAGCTTTAGGAATAATGCAAATCTAAGCCCAAGCT

22045 ATATTTTTCAGTCCTCACATTGTCTAAAAATCTTTCTAAATGGCTTTGTTATTGAATTTAT  
CTCATTTTATATCTGTGCCAACAGCATTTTCATCCTTTCTCTTCATAATTTCTTTTACAA  
ACAGCTGCTCAAGAGGAAGGCTCAAAGTCTCAAGGCTGAGCAGTAAATGACTTTTGTAG  
TACTAGATGAAAAGGGCTTTCCTGAGGAAATGAAAACCTAAAACATGAAAAGAAGATAAA  
CAGAATTTGGACAGTGAGATATAGAGCATATAATATTCTGCTTCTAAAGTAATATTCTTC  
[C, A, T]  
AGGAAAGTGAGGGCGTTTCCCTGGCTGTTAGGCCAGAAATCATATTCCTATATTTCTTT  
GATAGCTTTAGGAATAATGCAAATCTAAGCCCAAGCTTCAGAATAGACTAAGAAGTATT  
AGCTTAGCTGCCATGACAAAATACCATAGGCTGGATGCATTAAACAATGGAAATTTAGTT  
TTTCACAGTCTGGGAGCTGGGAAGTTTAAGATGAGAGTGCCAGCATGTTGGGTGTAG  
TGAGGGCTCTCTTTCTGGCTTGACAGATAGACCCCTTCTCACTGTATTGTCATATGGCAGA

22061 CATTGTCTAAAAATCTTTCTAAATGGCTTTGTTATTGAATTTATCTCATTTTATATCTGT  
GCCAACAGCATTTTCATCCTTTCTCTTCATAATTTCTTTTACAAACAGCTGCTCAAGAGG  
AAGGCTCAAAGTCTCAAGGCTGAGCAGTAAATGACTTTTGTAGTACTAGATGAGAAGGG  
CTTTCTGAGGAAATGAAAACCTAAAACATGAAAAGAAGATAAACAGAATTGGACAGTG  
AGATATAGAGCATATAATATTCTGCTTCTAAAGTAATATTCTTCTAGGAAAGTGAGGGCG  
[G, T]  
TTCCCTGGCTGTTAGGCCAGAAATCATATTCCTATATTTCTTTGATAGCTTTAGGAATA  
ATGCAAATCTAAGCCCAAGCTTCAGAATAGACTAAGAAGTATTAGCTTAGCTGCCATGA  
CAAAATACCATAGGCTGGATGCATTAAACAATGGAAATTTAGTTTTTACAGGTCTGGGA  
GCTGGGAAGTTTAAGATGAGAGTGCCAGCATGTTGGGTGTAGTGAGGGCTCTCTTTCT  
GGCTTGACAGATAGACCCCTTCTCACTGTATTGTCATATGGCAGAGAGAGAGAGAGAGA

22348 GAAAGTGAGGGCGTTTCCCTGGCTGTTAGGCCAGAAATCATATTCCTATATTTCTTTGA  
TAGCTTTAGGAATAATGCAAATCTAAGCCCAAGCTTCAGAATAGACTAAGAAGTATTAG  
CTTAGCTGCCATGACAAAATACCATAGGCTGGATGCATTAAACAATGGAAATTTAGTTTT  
TCACAGGTCTGGGAGCTGGGAAGTTTAAGATGAGAGTGCCAGCATGTTGGGTGTAGTG  
AGGGCTCTCTTTCTGGCTTGACAGATAGACCCCTTCTCACTGTATTGTCATATGGCAGAGA  
[-, A, G]  
AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGGGATCTTTCTCTTGCTTTCTATTATAAGG  
CCATAGTCTCTGTTGGATCAGGGTTCCATTCTTATGACTTTATTGACTTTACCCCTTAA  
GATGCTATCTCCAGATATAATCACACGGTGGGTAGGGCCTCAACATTTGGATTTGGGAG  
GGACAGAGCTCAGTCCATAGCAAAGGATAATGCAGAGGGTTGGATATTAAAGTAGCTA  
CACAATTTTAAATATAAATATTTTATGGTAACTTTTTTTTTTTTTTTGAGATGGAGTCTAG

FIGURE 3, page 33 of 42

22682 ATCTTTCTCTTGCTTTCTATTATAAGGCCATAGTCCTGTTGGATCAGGGTTCATTCTTA  
TGACTTTATTTGACTTTACCCCTAAGATGCTATCTCCAGATATAATCACACGGTGGGT  
TAGGGCTCAACATTTGGATTGGGAGGGACACAGCTCAGTCCATAGCAAAGGATAATGC  
AGAGGGTTGGATATTTAAAGTAGCTACACAATTTTAAATATAAATATTTATGGTAACT  
TTTTTTTTTTTGGAGATGGAGTCTAGCTCTGTTGCCAGGCTGGAGCGCAATGGTGCGA  
[A, G, T]  
CTCAGCTCACTGCAACCTCCGCCTCCAGGTTCAAGCAATTCTCCTGCCTCAGCCTCCTG  
AGTAGTTGGGACTATAGGCACGCGCCACCACGCTGGCTATTTTTTTTTTATTTTACTA  
GAGACGGGTTTGCACCATATTGGTCAGGCTTGTCTCGAACTCCTGACATCAGGTGATCCA  
CCCATCTTGGCCTCCCAAAGTGTGGGATTACAGAAGTGAGCCACCGCGCCTAGCCAGCA  
GCTTTACTGAGATGTAATTCACATGCCATAAATTCATTTTCTAAAGTATACAATTCAGT

22783 ATATAATCACACGGTGGGTTAGGGCTCAACATTTGGATTGGGAGGGACACAGCTCAGT  
CCATAGCAAAGGATAATGCAGAGGGTTGGATATTTAAAGTAGCTACACAATTTTAAATA  
TAAATATTTTATGGTAACTTTTTTTTTTTTTTGGAGATGGAGTCTAGCTCTGTTGCCAGG  
CTGGAGCGCAATGGTGCGATCTCAGCTCACTGCAACCTCCGCCTCCAGGTTCAAGCAAT  
TCTCCTGCCTCAGCCTCCTGAGTAGTTGGGACTATAGGCACGCGCCACCACGCTGGCTA  
[-, T]  
TTTTTTTTTATTTTACTAGAGACGGGTTTGCACCATATTGGTCAGGCTTGTCTCGAACT  
CCTGACATCAGGTGATCCACCATCTTGGCCTCCCAAAGTGTGGGATTACAGAAGTGAG  
CCACCGCGCTAGCCAGCAGCTTACTGAGATGTAATTCACATGCCATAAATTCATTTT  
CTAAAGTATACAATTCAGTGACTTAAACATTTATTTATTTTAAATTGACAGAATTACA  
TGTATTTATCATGTACAACATGATGTTTTGAAGTATATGTACATTGTGGAGTGACTAAGT

23448 TTCTCTTAGTATTTTCAAGAATATAATATATTATTATTAATTGTAGTCTTCATGTTGTA  
TAGTGGAGCTCTTGAACCTATTCCTCATGTCAAGCTGAAATTGTGTCTCTTTAACACAA  
ACCATACCCGACTCCCAAAGTATCTGCTCTCTGCTTCTATGAGATTAACTTTTCTGAT  
TCCACATGAGTGAGATCATGCAGTATTTATTTGTCTTACCTGGCTATTTTCATTCATAT  
TGTTACAGATAACAGGATTTCTTCTTTTTTAATGGCCGAATAGTTTTCTATTGTATAT  
[A, G]  
TATAGCACATTTTCTCTCTTCATGCATTGGTGGACACTTAGGTTGATTCCGTATCTTGGC  
TATCGTGAATAGTGCTATAATGAACATGGGAATGCACATGGCTCTTTGACATATTGATTT  
CATTTTATATATGTATATATATATGTATACACACACATACATACAGTGGTGGGATTGC  
AGGATCATATGGTAGTTCTATATTTAATTTTAAAGGAACTCCATACTGCTTTCATAAT  
GGCTGTATTAGTTTAACTCCTCACCAACAGGGTGCAAAAGTTCCCTTTTCTCTACATACT

24960 TTTGTTCTAGAGTATAGTTTAAAGTCTGATGTTTCTTACTGATTTTCTGTTGAGATGATTT  
GTCTATTGCTGAAGGTAGGGTGTGAAGTCCCCTACTATTGCTGTATTGCAGTCTCTCTC  
TCCTTTTCAACGATTAATGGTTTTTATTTTATTTTATTTGTTGTTGTTGTTGTTGTTGTT  
TGTTGTTTTTGGAGCGGAGTCTCACTCTGTCAACAGGCTGGAGTGCAGTGGCAGGGTCTC  
GGCTCACTGCAGCCCCGTCTCACGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTC  
[G, A]  
CTGGGACTACAGGCGCATGCCACCACGCCCAGCTAATTTTTGTATTTTATAGTAAAGACGG  
GGTTTCAACATGTTGGCCAGGATGGTCTTGATCTCTTGACTTCATGATCCACCGCCTTG  
GCCTCCCAAAGTGTGGGATTACAGGTGTGAGCCACCACCCCTGGCCAATGTTTGGTATT  
TATCTTTAGGTGCTCTGATGTTGGGTTTATATATATTTATAAAAAACAATAGCTACATAA  
CTTATTAAGGGATATGCAATATAAAATATATAAATTGTGACACTGAAAATTTAAATGGG

24983 TCTGATGTTTCTTACTGATTTTCTGTTGAGATGATTTGTCTATTGCTGAAGGTAGGGTGT  
TGAAGTCCCCTACTATTGCTGTATTGCAGTCTCTCTCCTTTTCAACGATTAATGGTT  
TTTTTTTTTATTTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTTGGAGCGAGTCTC  
ACTCTGTCAACAGGCTGGAGTGCAGTGGCAGGGTCTCGGCTCACTGCAGCCCCGTCTCA  
CGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTCTGTTGGGACTACAGGCGCATGCCA  
[T, C]  
CACGCCAGCTAATTTTTGTATTTTATAGTAAAGACGGGGTTTCAACATGTTGGCCAGGAT  
GGTCTTGATCTCTTGACTTCATGATCCACCGCCTTGGCCTCCCAAAGTGTGGGATTAC  
AGGTGTGAGCCACCACCCCTGGCCAATGTTTGGTATTTATCTTTAGGTGCTCTGATGTTG  
GGTTCAATATATTTATAAAAAACAATAGCTACATAACTTATTAAGGGATATGCAATATA  
AAATATATAAATTTGTGACACTGAAAATTTAAATGGGAGGAGTGGAGTAAAGTACCTTC

25390 AGTGCTGGGATTACAGGTGTGAGCCACCACCCCTGGCCAATGTTTGGTATTTATCTTTAG  
GTGCTCTGATGTTGGGTTTATATATATTTATAAAAAACAATAGCTACATAACTTATTAAG

FIGURE 3, page 34 of 42



GGATATGCAATATAAAATATATAAATTGTGACACTGAAAATTTAAAATGGGAGGAGTGGAGTAAAGTACCTTCATATAACTTACTATTATATCCTCTTATTGAATTGACCCCTTTATCATATATAGGAACCTTGTCTCTCTTACAACTTCTGACTTAAAGTTTGTCTTATATGATA  
[T, C]  
AAGTAAAGTTACTCCTGCTCTCCTTTGGTTTCTGTTTCCATGGAATATCTTTTCCATTCTTCACCATCAGTCTGTGTGATTTTTTACAGATGAAATGAGTCTGTCTGATGGGCAGCATATAGTTGGATCTAGTTTTTTAATCCACTCAGACACTGTGTTTTTTGATTGGATAAATTAATCCATTCTGTTCAAGGTAATTATTGATAAGTAAGGACTTTGTACTACCATTTTGCTTATTGTTTCATGGTTCTTTATAGATCCTTTATCTTTTCTCCTCTCTGCTGTCTTTTTTTT

26060 GGTTTTTGGTTTGTGGTTACCAAGAGGTTACAAAAACATCTTAAGAGTTATAATAGTTTATTTTAACTTGATAACTTAATTTTTATTGCAAAAACCCCCAAAAACAAAAAATCTACACTTTTACTTAATCCCTGAAATTTTGAATTTTGTATGTCACAGTTTACCTCTTTTCATATTGTGTATCCCTTAAATTATTGTAGCTATTATTACTTTTAAAGTTTTCTCTTCTCTACTACAGATGTAAGTGATTGCATACCATCATTACAGTATTATTTGAATTTACCTGTGTACTTT  
[C, T]  
TTTTATCAGCCAGTTTATACTTTTCTGATGTTTTTGTGTTACTCATTAGCATCTTTTCTTTCAGCTTGGAGCTCCTTTTACGTTTCTTATAAAATAGGTGCGGTCTGATTTATCTCCCTCAGCTATTGTTTGTCTGGGAAAGTATCTCTCCTTCATTTCTGAAGGACACTTTGCTGGGTACATTACCTTGGTTGGTATTTTTCTCCTTGAACGCTTTAAATATATCATCCCTTTCTCTCCTGACCTGTAGGTCTCTGCTGACCAGTCTGTTTCCAACCATATTGGGACTGTCTTA

30245 ATTTTAAACCATCCATTGTTTCTGCTTCTCTAGATAACCTGACTAATATATAATTGGTATGAAGTGATATCTCATGGCTTTGATTTATATTTCTTTCATGGCTAGTGACTTTTTTTGTACTTTTGGGATATTGTTATTATTATTATTATTACTAGTGTTTATACTTCTTCAGTAAAGTGTTAGAAACAATTTTTAAAGGCAGAATGTGACCAGAGTTTCCCTGTAGTTATATAACCATCATGGACCTTCCCTCAAGTGCTAAGCCATTAGTGTACTCATGTCACTCCAAATGTCAG  
[C, G]  
TTGTTTTCTTCCATTTCACTGTCTCTTTGTGTCCCAAACCTGAATTCATGGGAAAAACATCTGAATGGTGCTTAATATGGTTTGGATATTGTCCCCCTCCAAATCTCATGTTGAAATATGACCTCCAGTGTGGAAGTAGGACTACTTGGGTACGAGAGTGGATCCTTCATTAATGGCTTGGTAATAAGTGAAGTCTATTTAGTTTCTGAAAGCTGTTTGTGATAAGAGCCTGGCATCTCATTCTCTTGTCTCTCTCACCATCTGACACACTTGCTCACCTTTTTTCTTCAGCCA

33664 TTCCAGAGTGTAGAAGTACACTGTCCTATCCTTTCTAGGAGATCATTATAACACCAAAAGCAGACAGTATATGAAACAGGGAAATTAGAGGCCAAGATACCTATGACTTATATGTAAAAATTTAAAGAAATATAGCAAACTGAATCAGCCATTTTAAAAATATACCACAATCAATGCATTCAATAAGAGCAGCTTAACAAAATTTGTTAGAGGCATTAAAGAAGACTCAGTATAGAAAGATGTACCTTCTCTCCAAATTGGTGATAGAGATTCAATGCCATTAAAAAACCCACCT  
[G, T]  
GTTTTTTTGGAGAACTTGTCAAGCTGAGTCTCAAATTTATATCAAAGAGCAAAGGCCTAAATATCCAGGACATTCCTGAAGAACTGTAAAGAGCCAGGGGCTGCCCTATCAGATACCAGGGTTGTTATTAAGCCATAACCAAGTCAGTGTCTTCTACAGAAACAGACAAGTTAAACAGTGAAACATATAGAGAGCCAGAAACAGACCCATCCATTTTGGATTTGTACAGTGAAAGAAGTAGCTTTGCAAACTTTGGGAAAAGGAGAGTGTGTGCAATAGATGATGCTCG

33883 TAAAGAAGACTCAGTATAGAAAAGATGTACCTTCTCTCCAAATTGGTGATAGAGATTCAAATGCCATTAAAAAACCCACCTGGTTTTTTTTGAGGAACCTGTCAAGCTGAGTCTCAAATTTATATCAAAGAGCAAAGGCCTAAGAATATCCAGGACATTCCTGAAGAACTGTAAAGGAGCCAGGGCCTGCCCTATCAGATACCAGGGCCTGCCCTATCAGATACCAAGGGTTGTTATTAAGCCATAACCAAGTCAGTGTCTTCTACAGAAACAGACAAGTTAACAAGTGAAACATAATAGAGAGCCAGAAACAGACCCAT  
[C, A]  
CATATTTTGGATTTGTCTACGTGAAAGAAGTAGCTTTGCAAACTTTGGGAAAAGGAGAGTGTGTGCAATAGATGATGCTCGTGCTCATGCAGACAAAAGGAAATTTGGGATACCTGCCTCTTACCGTACACAAACCAACCTAAACGTGAAAGTTAACTATAACAGCTTGAGGTGGTGGGAAGAAATATCTTTATCTCAGTGTAGGGAAGAATTTATTTTAAAAAGAAGACACAAAAAGCCATACATAGGAATGAAAAGATTGAATTCAGCTGCATTAAAAAGATTAAATTCAGCTG

34373 TATCTTTATCTCAGTGTAGGGAAGAATTTATTTTAAAAAGAAGACACAAAAGGCCATACATAGGAATGAAAAGATTGAATTCAGCTGCATTAAAAAGATTAAATTCAGCTGCGTTAAAAATCAAGAGCATCTGTACTTGGACAGCATAGAGTGGAAAGACAAAGAGAAGGTATTTGCCAGCTTATAACTTGAAGGATTAGAATGAATGATATAAAGAAGTGTAAATAAGAAAAAGACATACAACCGGTTAGAAAAACGGGCAAGACATGAACAGCATATTTACCGTGAAGGAAAACAGC

FIGURE 3, page 35 of 42

[G, A]  
 GTAGCAAATGAACATGGTAAGAGATGCTCAACACGTTTAGTAATTTGAAGGGAAATGCAA  
 GTTATACCCACAGCAAGACTATCTTATCTAGGAAGTTTGTCAATACCCTAAATGTTCTGT  
 GGTTTTAAGCTACAGAGTTTGTAAATTCATTTATTTTCAATAAAATACTCAGTGGCAGGC  
 ACTGTTTTAGAAACCTTGGTTATAACTTTGAATGAAATAAAAAAATCCTTGCCTTGTG  
 GAGGATGCTTATGTGTGGGGAGTTGGGTGGTGGGGTCAAACAACAATTACATTAATAATAG

34558 ACTTGAAGGATTAGAATGAATGATATAAAGAACTATGTAAATAAGAAAAAGACATACAAC  
 CGGTTAGAAAAACGGGCAAAGACATGAACAGCATATTTACGTGAAGGAAACAGCGGTAG  
 CAAATGAACATGGTAAGAGATGCTCAACACGTTTAGTAATTTGAAGGGAAATGCAAGTTA  
 TACCCACAGCAAGACTATCTTATCTAGGAAGTTTGTCAATACCCTAAATGTTCTGTGGTT  
 TTAAGCTACAGAGTTTGTAAATTCATTTATTTATTCAATAAAATACTCAGTGGCAGGCACTG  
 [G, T]  
 TTTAGAAACCTTGGTTATAACTTTGAATGAAATAAAAAAATCCTTGCCTTGTGGAGGA  
 TGCTTATGTGTGGGGAGTTGGGTGGTGGGGTCAAACAACAATTACATTAATAATAGAAAT  
 AGTGACATAAATAAACCTATAAATATTGCAACCCAGAGTTATATTATAAATGTAAGTAGT  
 GACTAGGACTCTCATGCAGATATACCTCTGTGCTGGGACAAATGAAAGTTTAAGTGTAAT  
 TTCCCATATGCAAGTCAAATAAAAAAGTGACACTAGAAAACACAATAATGAATATCTGAA

43929 GGCATTTAAGTATTCTGCCATAGGGAAGTGTAAGTTGTAGGCTTTACTTTTTATAGG  
 TACTATATTGTCCAAATAATCTCAGCACCTCATGGTTGCTAAGGATCTGTGTCCTTGTTT  
 GGTGAGATTATGTTTATCTCTGGCATAAGGCACTTAACAATATTTCATTAAAGGTTACAGA  
 ATCTTTTTTGCTTCATCTGCTTAGCATTTTCATACCAGTTTGTTCACCAAACCTTTCAA  
 TTTTGATTGTTTCATTAATATTTCTGCATACTGATGTAAACCAAGTTCTATTATTGTGCAA  
 [T, A]  
 CTGCTCTGAAACCTTAGGAACTCTCTGAAGGAGTTTATTTATTTTTGTTTTGTTTT  
 TTGTTTTGTTTTGTTTTTTTTGAGACGGAGTCTTGCTCTGTTGCCAGGCTAGAGTGCAG  
 TGGTGCGATCTCGGCTCTCTGCAAACTCGGCCTCCGGGGTTCACGCCATTCTCCTGCCTC  
 AGCCACCGGAGTAGCTGGGACTACAGGCACCCACCCTGCGCCTGGCTAATTTTTTTTTGT  
 ATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTT

44309 TTGAGACGGAGTCTTGCTCTGTTGCCAGGCTAGAGTGAGTGCGATCTCGGCTCTC  
 TGCAAACTCGGCCCTCCGGGGTTCACGCCATTCTCCTGCCTCAGCCACCGGAGTAGCTGGG  
 ACTACAGGCACCCACCCTGCGCCTGGCTAATTTTTTTTTGTATTTTAGTAGAGACGGGG  
 TTTACCGTGTTAGCCAGGATGCTCTCGATCTCCTGACCTTGTAAATCCGCCCGCCTCGCC  
 TCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGTGCCCGGCCTTTTTTTTTTTTTTTT  
 [T, -, C]  
 TTTATGGGCTTGCTCTTCTACACTTCAGATTGACTAAATTAATATGCATTAAATGAAGT  
 CAGGAGTTCACATTGCCACTAGTAACAATGCCTAAGCTTACATAAAGCATTATAAAATTG  
 TTGGTGATTAGTGCCTTCTCAGCTATGAGTATAAGATAATATTATACTAGTAGTTCAGTT  
 GCCTAGATAAATTGTACACTATGTGAAGTTTATTTACATAATTCTTACGGTATTTTTTA  
 AGGTAGTTGATAACAGTTGAGACTACAATTGTATCTCCATTTATTGATAGTAAATGAA

44997 GAATTGTAAAAATATTATTATAGAATTGTTTCTCTCAAACTATAGTAATGTAGAATAGGT  
 TGAAGGGGTGATGATTGAAACAATACCTCTCCATTAGCTAAATTTATATAGAATCTAT  
 TGCATGTTTTAAATGATAAGTCAGATTTATAAAAAATATTTTTATAAACAGTAGGAAATGA  
 GTTTAGGGGTATTCACATACAGTTTAAATTTTATTTACATATTTAAACATATCATGGT  
 ATAAATATGATGTGGATATAAATTTGAGATAAAGGAAGTATTGTTTAAGAATTGATGAAC  
 [T, G]  
 AATTTCTTAAAGATGTCATCACCAGTTGGTTTTCTAGCCTTATGAAAAATGGTTGCAAT  
 AAAAAAGATTGACTATGATAAAATGCTGCCCTTTCATTTAACCTAGACCAAGAGAAAAC  
 ATACTGTGAATCTATGATGAATGAAAGAAAGTTGTAAGTGTGGTTTGTATATTGTAA  
 TTAGTGTATTTTTCATTTCTTGTGAAGTACTGACTTTGTTTATTGTGAGTAGACA  
 ACTTATAATCTATGTACTCAAATTTGGTTTAGTATAAATCTAGGGAATGAAGTTCATATT

46538 TGTATACTTATGGTCAACACTTTTTATATTGTCTGTAGATTTCTGTACAAAAAGATTC  
 TGACACTGTTTTAAGCCAGCATTCCTTCAGAAATGTACCCAAATCTCAAATTTATTTAGG  
 GGCAAAGCTAATGCTTTAAAGAAAAAGGAGA  
 [A, G]  
 GGGATTGGTGTGTGTTTTCTTTAGGAACAGTAGTAAGTTGACTTTTAGAGAACTTGAAT  
 AAGCATTTATTTTTCTTTGTCCTATTTTATTGTGAAGTTTATTTATTTAAAAATAAAT  
 GGATTTCTCTGGAATTTAGTTTCTGCAAATTTGAGGAGTTTCCAAAGTCAACCTTCAGGT  
 TTGATACTTCTCTAGAAAGACTCACATAACTCACTGAAAGCTTATTACCCCTGGTTATGG

FIGURE 3, page 36 of 42

TTTATTACGGGGAAAAGATGCGGATGAAAATCAGTCAAGTAAAGAAGCACATAGGGCAGA

48153 TTATATCATTCTGCTTTTATTTTATAGGTTACGGTTCAAAATCAGACAAAATGAACATAT  
TTGGTGGCTTTTCGACAGATGGTAAAAGAAGGAGGTATCCGCTCGCTTTGGAGGGGAAATG  
GTACAAAAGTCATCAAATTTGCTCCTGAGACAGCTGTTAAATTTCTGGGCATATGAACAGG  
TAATTTGTTATCACCGTGGAATTTATTAACAAAGAGGAGTTAGTAAACGGATTCAATAAA  
TGTTAATGTATAATGCTTTTGGGATTCTTGTTTAAATACATGATAATCTTTCACATATAC  
[T,C]  
CCATAAGGAGGATCACTTATAGGAGATTAGACTAAATAAAATCAGAGATTTCTCATGACC  
AAGTTATGGGATTCTTAATTCATCATATTATTTATAAAGTTTTTTTTTTCTAAGTAGTTC  
TTAAAGGAAGGGTAGAATTTTAGTTTATTCATTCTGAATCCTGAGCAGAAGCAGCACACT  
AACATAAGTTTTTATGAAAGTGTCACAATCTAACCTCTGGAAGGAAAACATAAGTTGAAG  
TCCTTTGTGTAATTTGACGTTGCTGTAAAATTGAGCTGAGTTTGGAGTGACACCTCCATG

48288 AAATTGCTCCTGAGACAGCTGTTAAATTTCTGGGCATATGAACAGGTAATTGTTATCACCC  
GTGGAATTTATTAACAAAGAGGAGTTAGTAAACGGATTCAATAAATGTTAATGTATAATG  
CTTTTGGGATTCTTGTTTTAATACATGATAATCTTTCACATATACCCCATAGGAGGATC  
ACTTATAGGAGATTAGACTAAATAAAATCAGAGATTTCTCATGACCAAGTTATGGGATTCT  
TTAATTCATCATATTATTTATAAAGTTTTTTTTTTCTAAGTAGTTCTTAAAGGAAGGGTA  
[G,T]  
AATTTTAGTTTATTCATTCTGAATCCTGAGCAGAAGCAGCACACTAACATAAGTTTTATG  
AAAGTGTCACAATCTAACCTCTGGAAGGAAAACATAAGTTGAAGTCCTTTGTGTAATTT  
GACGTTGCTGTAAAATTGAGCTGAGTTTGGAGTGACACCTCCATGAAGGCAGGGGCGTGG  
CTTCTCCCATGTACTCCAGCACCTAGACAGAGCTTGGCATGTGATAAGTTTCAAGCGA  
GTGTTGAATGAGTCAATGAATGAACAAATGCATTTACCTCTGAATCACTTCTCTGTCGGC

48412 TGGGATTCTTGTTTTAATACATGATAATCTTTCACATATACCCCATAGGAGGATCACTT  
ATAGGAGATTAGACTAAATAAAATCAGAGATTTCTCATGACCAAGTTATGGGATTCTTAA  
TTCATCATATTATTTATAAAGTTTTTTTTTTCTAAGTAGTTCTTAAAGGAAGGGTAGAAT  
TTTAGTTTATTCTGAATCCTGAGCAGAAGCAGCACACTAACATAAGTTTTATGAAA  
GTGTCACAAATCTAACCTCTGGAAGGAAAACATAAGTTGAAGTCCTTTGTGTAATTTGAC  
[G,A]  
TTGCTGTAAAATTGAGCTGAGTTTGGAGTGACACCTCCATGAAGGCAGGGGCGTGGCTTC  
TTCCCATGTACTCCAGCACCTAGACAGAGCTTGGCATGTGATAAGTTTCAAGCGAGTGT  
TGAATGAGTCAATGAATGAACAAATGCATTTACCTCTGAATCACTTCTCTGTCGGCTTTT  
GTTAACTTGGATTATTTGAGCTATTGCTTCAGCCTAACTCAATGTAAAGGGGAAATACAG  
AGGTAAGTTTTAGAGTTTGGGTTCTCTTTATGGTCATTAGCAGAACTGTCTAGTTGAGCA

48446 CATATACCCCATAGGAGGATCACTTATAGGAGATTAGACTAAATAAAATCAGAGATTTCT  
TCATGACCAAGTTATGGGATTCTTAATTCATCATATTATTTATAAAGTTTTTTTTTTCTA  
AGTAGTTCTTAAAGGAAGGGTAGAATTTTAGTTTATTCTGAATCCTGAGCAGAAGC  
AGCACACTAACATAAGTTTATGAAAGTGTCACAATCTAACCTCTGGAAGGAAAACATA  
AGTTGAAGTCCTTTGTGTAATTTGACGTTGCTGTAAAATTGAGCTGAGTTTGGAGTGACA  
[C,G]  
CTCCATGAAGGCAGGGGCGTGGCTTCTTCCCATGTACTCCAGCACCTAGACAGAGCTTG  
GCATGTGATAAGTTTCAAGCGAGTGTGTAATGAGTCAATGAATGAACAAATGCATTTACC  
TCTGAATCACTTCTGTCGGCTTTTGTAACTTGGATTATTTGAGCTATTGCTTCAGCC  
TAACTCAATGTAAAGGGGAAATACAGAGGTAAGTTTTAGAGTTTGGGTTCTCTTTATGGT  
CATTAGCAGAACTGTCTAGTTGAGCAGCCACAGATTATGTTTTCCATTATTTATTCATC

48456 ATAAGGAGGATCACTTATAGGAGATTAGACTAAATAAAATCAGAGATTTCTCATGACCAA  
GTTATGGGATTCTTAATTCATCATATTATTTATAAAGTTTTTTTTTTCTAAGTAGTTCTT  
AAAGGAAGGGTAGAATTTTAGTTTATTCTGAATCCTGAGCAGAAGCAGCACACTAA  
CATAAGTTTTATGAAAGTGTCACAATCTAACCTCTGGAAGGAAAACATAAGTTGAAGTC  
CTTTGTGTAATTTGACGTTGCTGTAAAATTGAGCTGAGTTTGGAGTGACACCTCCATGAA  
[G,C]  
GCAGGGGCGTGGCTTCTTCCCATGTACTCCAGCACCTAGACAGAGCTTGGCATGTGATA  
AGTTTCAAGCGAGTGTGTAATGAGTCAATGAATGAACAAATGCATTTACCTCTGAATCAC  
TTCTCTGTCGGCTTTTGTAACTTGGATTATTTGAGCTATTGCTTCAGCCTAACTCAATG  
TAAAGGGGAAATACAGAGGTAAGTTTTAGAGTTTGGGTTCTCTTTATGGTCATTAGCAGA  
ACTGTCTAGTTGAGCAGCCACAGATTATGTTTTCCATTATTTATTCATCATTGTTTATC

48789 GCACCTAGACAGAGCTTGGCATGTGATAAGTTTCAAGCGAGTGTGTAATGAGTCAATGAA

FIGURE 3, page 37 of 42

TGAACAAATGCATTTACCTCTGAATCACTTCTCTGTCGGCTTTTGTAACTTGGATTATT  
 TGAGCTATTGCTTCAGCCTAACTCAATGTAAAGGGGAAATACAGAGGTAAGTTTGTAGAGT  
 TTGGGTTCTCTTTATGGTCATTAGCAGAACTGTCTAGTTGAGCAGCCACAGATTATGTTT  
 TCCATTATTTATCCATCATTTGTTTATCAAGGACTGTAAGGGCCTTGAAATCAACTCCC  
 [C, -]  
 CCCCCATAGTTTTTGTATTATTCATGTAGATTTTAGATTATTCTGGAGAGTGTTTTGT  
 CTTGAGCAACAGAATACTCTTGAGAAGATTACGAAGTCCAGTGGTATCCTTTTCTTGCC  
 TAGGAAATAGAGAAGCAAAAAAAAAAAAAAAAAAAATTAAAGAAAATCTAGTCTCCAGG  
 ATTTTAATTAGAACCTATCCTTGGAAGGCTATTTTCTTATATGAAGGTTTGAAGATT  
 AAATCATGATTATTAAGGGCTAATGTTTGAGATACCCTTAGGTTATTCTGACCACATACT

48859 CATTACCTCTGAATCACTTCTCTGTCGGCTTTTGTAACTTGGATTATTTGAGCTATTG  
 CTTAGCCTAACTCAATGTAAAGGGGAAATACAGAGGTAAGTTTGTAGAGTTTGGGTTCTC  
 TTTATGGTCATTAGCAGAACTGTCTAGTTGAGCAGCCACAGATTATGTTTCCATTATTT  
 ATTCATCATTTGTTTATCAAGGACTGTAAGGGCCTTGAAATCAACTCCCCCCCCATAG  
 TTTTTGTATTATTCATGTAGATTTTAGATTATTCTGGAGAGTGTTTGTCTTGAGCAA  
 [G, C]  
 AGAATACTCTTGAGAAGATTACGAAGTCCAGTGGTATCCTTTTCTTGCCCTAGGAAATAG  
 AGAAGCAAAAAAAAAAAAAAAAAAAATTAAAGAAAATCTAGTCTCCAGGATTTTAATTA  
 GAACCTATCCTTGGAAGGCTATTTTCTTATATGAAGGTTTGAAGATTCAAATCATGAT  
 TATTAAGGGCTAATGTTTGAGATACCCTTAGGTTATTCTGACCACATACTTGGATTTTAT  
 GATAGGAAAGCCACAGCCTAAAATAAATAAATACTCAATGCAGTTATTTTCAAGTATGCAAG

49126 GATTATTCTGGAGAGTGTTTTGTCTTGAGCAACAGAATACTCTTGAGAAGATTACGAAG  
 TCCAGTGGTATCCTTTTCTTTGCTAGGAAATAGAGAAGCAAAAAAAAAAAAAAAAAAAAA  
 ATTAAGAAAATCTAGTCTCCAGGATTTTAAATAGAACCTATCCTTGGAAGGCTATTTT  
 CCTTATATGAAGGTTTGAAGATTCAAATCATGATTATTAAGGGCTAATGTTTGAGATACC  
 CTTAGGTTATTCTGACCACATACTTGGATTTTATGATAGGAAAGCCACAGCCTAAAATAA  
 [A, G]  
 TAAATACTCAATGCAGTTATTTTCAAGTATGCAAGAAGTTTGGTATTTTGAAGAAAGTCCAT  
 GGGTATTGCAAGCAAAATATGCACATTTTGTCTTATGCCATTTGTGAGATTCTTACCTTGG  
 ATACCACCAACAGGCATCCTCTGCTTCTGTCCACCCAAGCTCCTTCTGAGACCTCTTTA  
 TAGTATTGTGATTTCTGCACACTAACTTTCTTAGACATGAAGAGAAAGCTGTCTACACAG  
 TGTGGTGTAGTTTTCTTATGGGCTCTGGACCTATGGTGTGTTTTCTCTCCTCCTGCTGA

49378 TGACCACATACTTGGATTTTATGATAGGAAAGCCACAGCCTAAAATAAATAAATACTCAA  
 TGCAGTTATTTTCAAGTATGCAAGAAGTTTGGTATTTTGAAGAAAGTCCATGGGTATTGCAA  
 GCAAAATATGCACATTTGCTTTATGCCATTTGTGAGATTCTTACCTTGGATACCACCAAC  
 AGGCATCCTCTGCTTCTGTCCACCCAAGCTCCTTCTGAGACCTCTTTATAGTATTGTGA  
 TTTCTGCACACTAACTTTCTTAGACATGAAGAGAAAGCTGTCTACACAGTGTGGTGTAGT  
 [T, G]  
 TTCTTATGGGCTCTGGACCTATGGTGTGTTTTCTCTCCTCCTGCTGAAGGTCCATTCAT  
 CCCTCGGGGCTCTCTAAAAGCCACCTTCTGTGACAAGCATATACTAAGCATCTCAATCA  
 AAGCCAGTTCTCCCTGCTCCAGCCTCCCTCGAGTGCTGAATTGCAGAATATCCATTTT  
 TCATTGGATGATGGAAACCCATTGTTTTCCAGTGGATTGTAAATTACTTCGGGGTAAA  
 TAGGCTGTATATATTCTCAAATTTCCAGAGTATGTAAGTACTAGTCACTTTTAGATTGAGA

49482 TCCATGGGTATTGCAAGCAAAATATGCACATTTTGTCTTTATGCCATTTGTGAGATTCTTAC  
 CTTGGATACCACCAACAGGCATCCTCTGCTTCTGTCCACCCAAGCTCCTTCTGAGACCT  
 CTTTATAGTATTGTGATTTCTGCACACTAACTTTCTTAGACATGAAGAGAAAGCTGTCTA  
 CACAGTGTGGTGTAGTTTTCTTATGGGCTCTGGACCTATGGTGTGTTTTCTCTCCTCCT  
 GCTGAAGGTCCATTCATCCCTCGGGGCTCTCTAAAAGCCACCTTCTGTGACAAGCATAT  
 [A, C]  
 CTAAGCATCTCAATCAAAGCCAGTTCTCCCTGTCCAGCCTCCCTCGAGTGCTGAATTG  
 CAGAATATCCCATTTTTCATTGGATGATGGAAACCCATTGTTTTCCAGTGGATTGTAA  
 ATTACTTCGGGGTAAATAGGCTGTATATATTCTCAAATTTCCAGAGTATGTAAGTACTAGG  
 CACTTTTAGATTGAGATAGATTTGTTTCTTGAATAGCTAGTACTTTAGGAAACTAAGAA  
 AAAGATCTTTTCAACCTGGTATGTAGCTCTGTCAAACACATCATCAGTATGGGGTAAACC

49741 CTCGGGGCTCTCTAAAAGCCACCTTCTGTGACAAGCATATACTAAGCATCTCAATCAAA  
 GCCAGTTCTCTCCCTGTCCAGCCTCCCTCGAGTGCTGAATTGCAGAATATCCCATTTTCT  
 ATTTGATGATGGAAACCCATTGTTTTCCAGTGGATTGTAAATTACTTCGGGGTAAATA  
 GGCTGTATATATTCTCAAATTTCCAGAGTATGTAAGTACTAGTCACTTTTAGATTGAGATA

FIGURE 3, page 38 of 42

GATTTTGTTCCTTGAATAGCTAGTACTTTAGGAACTAAGAAAAAGATCTTTTCAACCTG  
[G,A]  
TATGTAGCTCTGTCAAACACATCATCAGTATGGGGTAAACCTGTGTTCTCTGTGGGTGT  
CATTACCATAGTAGTGTCAATTGTATCATTGACAGTGTAAATAGTGTGGGGTAGTGTCTTG  
TGGTTTCAGCTGCCACTCTGTACTGACTGCTTTCCACTCCA

49840 ATCTTTTCAACCTGGTATGTAGCTCTGTCAAACACATCATCAGTATGGGGTAAACCTGTG  
TTCTCTGTGGGTGTGATTACCATAGTAGTGTCAATTGTATCATTGACAGTGTG  
[A,G]  
TAGTGTGGGGTAGTGTCTTGTGGTTTCAGCTGCCACTCTGTACTGACTGCTTTCCACTC  
CAACATCTTCTCTTTATCTCAACACTGTAGGTCTACCTGTGTACTGTGTGTTTCAGCAT  
CTGCTCTCCACTGACCCAGGAGTGCCTCCACTCAATATGGCCACCATGCATGGTCACTCT  
TTCTGCTACTCCCTGTCTCTGACCTGCTCCAGCAACACAGACAGACACCTTCTCTCT  
TCTATATGTCATATGGTGGGGAATGCCCTTTAGTACTTACTCAGGAGTTAGTTCTCTG

50102 CATTACCATAGTAGTGTCAATTGTATCATTGACAGTGTAAATAGTGTGGGGTAGTGTCTTG  
TGGTTTCAGCTGCCACTCTGTACTGACTGCTTTCCACTCCAACATCTTCTCTTTATCTC  
AACACTGTAGGTCTACCTGTGTACTGTGTGTTTCAGCATCTCTGCTTGCATGACCCAGGA  
GTGCTCTCCACTCAATATGGCCACCATGCATGGTCACTTTCTGCTACTCCCTGTCTCTCT  
GACCTGCTCCAGCAACACAGACAGACACCTTCTCTTTCTATATGTCATATGGTGGGG  
[G,A]  
ATGCCCTTTAGTACTTACTCAGGAGTTAGTTCTCTGGGAAGCCTTCTGTTCTAGTTTCC  
TTTTGTTACAGCACTTTCACATTGAATTCGACGTTCTCTGTACTTATCTGCTTTGTGAG  
ACTGTGAGCTTCTTTAGGCAGTAGCTACTTGTATTCTTAGCACCTTGCCAGTGCCAGGA  
AACCTTATTAAGTAAATGAAAAGACAGAACTGACAGACTGGAATTAGAGCTCAAGCTTG  
CCTCAATCTCAAGCCATTAAGATGAAGGGGAGCCGGGCGTGGTGGCTCACGCCCTAATC

50109 ATAGTAGTGTCAATTGTATCATTGACAGTGTAAATAGTGTGGGGTAGTGTCTTGTGTTTC  
AGCTGCCACTCTGTACTGACTGCTTTCCACTCCAACATCTTCTCTTTATCTCAACACTG  
TAGGTCTACCTGTGTACTGTGTGTTTCAGCATCTCTGCTTGCATGACCCAGGAGTGCCTC  
CCACTCAATATGGCCACCATGCATGGTCACTTTCTGCTACTCCCTGTCTCTGACCTG  
CTCCAGCAACACAGACAGACACCTTCTCTTTCTATATGTCATATGGTGGGGAATGCC  
[C,G,T]  
TTAGTACTTACTCAGGAGTTAGTTCTCTGGGAAGCCTTCTGTTCTAGTTTCTTTTGT  
ACAGCACTTTCACATTGAATTCGACGTTCTCTGTACTTATCTGCTTTGTGAGACTGTGA  
GCTTCTCTTAGGCAGTAGCTACTTGTATTCTTAGCACCTTGCCAGTGCCAGGAACCTT  
ATTAAGTAAATGAAAAGACAGAACTGACAGACTGGAATTAGAGCTCAAGCTTGCTCAAT  
CTCAAGCCATTAAGATGAAGGGGAGCCGGGCGTGGTGGCTCACGCCCTAATCCCAGCAC

50747 CCAGCCTGGGCAACGTGGCAAAACCCCATTTCTACAAAAATATAAAAAATTAGTTGGACG  
TGGGGGTGTGTGCCTGTAAGTCTGAGGATGCTGAGGTGGGAGGATCACTTGAGCTCGAGAGGC  
AGAGGTTGCAGTGAGCTGGGATCACACCATTGCAATCTAGCCTGGGTGATAGAATGAGAC  
CTTGTCTCAAAAAAATAAATAAATAAATAAAGGGGAAGATAAGGATTGGAACAGAA  
GGAGCAGCATGTGGACAGAAATGTAGGCACAAGAAGGCATCACTCACTGAAGAGACTGAA  
[G,A]  
GTGGTTCACTGTGCCTCAAGACTGGTGGAGTGTGTTCCGGAAGATAATGATGAAAGAG  
CTGGACAGATAAACAGGGGCCAATGTAATAGGAGTCTGGATTTTATTCTGAATATGGTA  
GGGGCTATTGTAGCATCTTATATAGGGAAGTGAATGAGTACATTACATTTAAGGAATA  
TCAACCTGAAAAAAGAGTGGAGACATTGTTGGGGGAGACTGAGGTAGACTAGAGGCAGGG  
AGAATATTTAAATAATTGAGGTAAGAAATGATGAACACCAGTATAAGGTGATGTCTTTAA

51272 TAGACTAGAGGCAGGGAGAATATTTAAATAATTGAGGTAAGAAATGATGAACACCAGTAT  
AAGGTGATGTCTTTAAGGAATGGAGAAGGGAATGAAGTGAAGAAATATTTGGAAGTAGAA  
TCAACAGAACTCACTGACTGACTGGATATGGAGGTGAGAAAGAGAAGAGTCAAGAATGAT  
ATTCTAATTTCTAACTTGAGTGAATGCAATCAAGAGAATACAATATCAGGTTCCATTTT  
GTGCATGCTGAGTTTGAGATGTGTGGGACATGTACAGGAGCTGTCCAGTAAGCAATTGG  
[G,A]  
TATATCAGCTAGCCATTAAGAGAGAGATCTTTGATAGAGAGGTTGTTGCTGAGTTGAGCC  
ATTGGAATGGGCAGGATCACTCAAGAAGAGCTTATAAATGAGAAGAAATCTAGGAATAAG  
TCCAAAGGGAGAAGTAAAAGAAGAACTTGCAAAGGACACTGAGAAGAAATAGCTCGAGG  
GATGGGAGAAAAATCCAGAGAGAGGGATGGCATAGGAGTCAGTGAAGGAAACGGTTTCAT  
GGGGGTCAGTACTACTGGGTAGTGAATATAATAAGAATATCTTTTAGGATTTCTCAACCC

52842 TCAGGGTGGTTTTGAGGGCTCAGTTAAGTCTCCTTTAGGAAGGTTGAGTTCTGTAGCCTT  
GGCAAGTTACTTAAAGTCTCTGTGACTATTACCTCATCTCTAAGATGGGGACTAAGCTTG  
GTGACATAGTTTTACATACCAGGCACAGTGCCTGACTTTTTGGCTCTGTCTGAACTCTT  
CCCTTTGTATATGGTATGTTTTCGGGGAATAGGAGCCTCAAGCACTTATCCTTTAAATATT  
TATCCTCCATCAGTCACTAAACGTTTACTCTGTACTTTTGATAGGTGCTGTGGGGTCCA  
[G, A]  
GGTATAAAAGGTACCTTCAAAGTTACTGTTAAAGTGCAGGAAGGTTTTAAGCAAATTAT  
GTTTAAATGATTTTGACAATCTGACATGCAGGAAAATTAATAGGGCCTATGCAGAAGAGGA  
GTTTTATGTAACACTCTGTAGTTTCAAGAAACAGAGCCCTTGAAGCAGTGATCTCTCTGG  
GGAGGAATGTCTGGTATTTGGGAATCTCATGAAATGATAATATACTTAATTTTTATCATG  
AGCAGCAAAACACAGATTTGCTAGGAGAAAGTCATCGTATGTTGTTGCATTGGGCACTTT

61837 GAGGAACCTCCATGTCAATTTCCATAGTAAGTACCTTTTTGTTTTTAAACATTTCTAT  
CAATGTACACCAAGATTCCAATTTCTCCATGTCTCCCCAACACCATTAAGTGGGGTGGT  
GGTCTACTACTATTGCTGTGTTGCTGTTTATTCTCCCTTCAGTTCTGTAAGTGTGTTGCT  
TCATATATTTAGGAGCTTAATATTAGGTCCATATGAAGTTATAATTTCTTCTCGGTAAAG  
TGACCCATTTATCATTATGTAATGTCCATCTTTGTCTCTGTGACAGTTTGTGTCTTAA  
[A, G]  
TCTATTTTGTCTGATGTAATTATGGCCACCCCTTTCTCTTTGGGTCCCGTTTTATGG  
AATATCTTTTTCCATCCTTTTCACTTTCAGCTTATGTGTGTCCTTAGATCTAAAGTGAGTC  
TCATAGATAAGGTATAGTTGATTCTGTATGTGTTATTCAGTCAAGTTTATATCTTTTA  
GTTAGGGGATTTAATCCATTTACATTTAAAGCAGTTACTGATAGGGAAGGACTTACTGTT  
GTCATTTGGCTAGCTACCTTTTTATCTTTGTCTGTGGCTTTTCTGTTTTTCCCTTCCTC

62018 CATATATTTAGGAGCTTAATATTAGGTCCATATGAAGTTATAATTTCTTCTCGGTAAAGT  
GACCCATTTATCATTTATGTAATGTCCATCTTTGTCTCTGTGACAGTTTGTGTCTTAAAA  
TCTATTTTGTCTGATGTAATTATGGCCACCCCTTTCTCTTTGGGTCCCGTTTTATGG  
AATATCTTTTTCCATCCTTTTCACTTTCAGCTTATGTGTGTCCTTAGATCTAAAGTGAGTC  
TCATAGATAAGGTATAGTTGATTCTGTATGTGTTATTCAGTCAAGTTTATATCTTTTA  
[A, G]  
TTAGGGGATTTAATCCATTTACATTTAAAGCAGTTACTGATAGGGAAGGACTTACTGTTG  
TCATTTGGCTAGCTACCTTTTTATCTTTGTCTGTGGCTTTTCTGTTTTTCCCTTCCTCT  
CTTCTGGCTTCTTCTGTGTTTTGTTGATTTTTTTTTTTTTTTGTAGTGATATGTTCTGAT  
TCCCTTCTCATTTCCCTTTGTGTGCATTCTATAGATGCTATTTTTGTGGTTACCATTGCA  
ACTACATAAAGCATATAAAGTTATAGCAACTTATTTTAAGCTGTTTACAACCTAACTTC

65562 GACTGAAATTCAGACACATGCAGTCTGATTCTAACCCCTCCTGTCTGCCAGCTCTGATCCA  
GAACTTTGCATGACTGATACGGCTGATAGATTGTCTATGGCTGATAGACTGTCATTTCTG  
ACCTAAAAGTCTGATCATTTTACATCTGTTTCAAGACATCTTTCAGCCCTTTCGGTGTGAGT  
TCCAAAGTTGTAGTGGGAATTTCAAAGCCTTTAATAATCTAGCCCACTTTGTTCACTC  
TCTGTGTAATAACCACATACAACAATTGGCTGCATCTCCATAGCACATGGTACTCCTCCC  
[A, G]  
TTGTCTTGGTTGTGCCAGCAACACTGGTTTTTCGCTTTCTCTTCTGCTGTTGAGGTCAT  
TTCCAAGGCCAGGTCTTTGTGCTTTTTCCCAAGCTTCCCAGAGCTTCTTCCATACTCCC  
CTTACTTCTGAGATTTAACTGTTCTCTTTCAGCGCTTGTCTAGTAAGAAGGAGGCAGC  
AGCAGCACTGTGGGGTGGTGGAAAGTGTACCAGCTTTGGAGTCAGACCATTGGATCTCAG  
CCCTACCATTTTCTACTTAGATTTTTTTTAGGACAAATTTCTCCATCTTTCTAAGCCTCCA

65780 TCTAGCCCCACTTTGTTCACTCTCTGTGTAATAACCACATACAACAATTGGCTGCATCTC  
CATAGCACATGGTACTCCTCCCGTTGTCTTGGTTGTGCCAGCAACACTGGTTTTTCGCTTT  
CTCTTCTGCTTGTGAGGTCAATTTCCAAGGCCAGGTCTTTGTGCTTTTTCCCAAGCTT  
CCCAGAGCTTCTTCCATACTCCCCTTACTTCTGAGATTTAACTGTTCTCTCTTCTCAGCGC  
TTGTCTAGTAAGAAGGAGGCAGCAGCAGCACTGTGGGGTGGTGGAAAGTGTACCAGCTTT  
[G, A]  
GAGTCAGACCATTGGATCTCAGCCCTACCATTTTCTACTTAGATTTTTTTAGGACAAATT  
TCTCCATCTTTCTAAGCCTCCAATTGCTCACTTACAAAATTGATATAACATTTACCTTGC  
AAGATTGGTATGGAAGGTAATTAACCCAGTATTTAGAACATAGTAATTAATAAATAACTA  
TTATTACCATCATTACTATAGTTAGGACACTCACTGTTAGGTGCTATACAAAGAGGATCA  
TAAAGGGATGTTGTCTTGGGCTTCTTGGAAATAAATGTTGTCTTTTACTGTATTTTGA

66092 TTGGATCTCAGCCCTACCATTTTCTACTTAGATTTTTTTTAGGACAAATTTCTCCATCTTT  
CTAAGCCTCCAATTGCTCACTTACAAAATTGATATAACATTTACCTTGCAAGATTGGTAT  
GGAAGGTAATTAACCCAGTATTTAGAACATAGTAATTAATAAATAACTATTATTACCATC

FIGURE 3, page 40 of 42

ATTACTATAGTTAGGACACTCACTGTTAGGTGCTATACAAAGAGGATCATAAAAGGGATG  
TTGTCTTGGGCTTCTTGAATAAATGTTGTCCTTTTACTGTATTTTAGAATATCATTCTG  
[G, A]  
GTCATAATTGTTTGTGTCATAATAATGAAACATACTTGAATATTAAATTACCCTCTTTT  
TTTATTTTTTAGCCATGTTAGAGGTTCCCCACAGCTGAATATGGTTGGCCTCTTTCGAC  
GAATTATTTCCAAAGAGGAATACAGGACTTTACAGAGGCATCACCCCAAACCTTCATGA  
AGGTGCTCCCTGCTGTAGGCATCAGTTATGTGGTTTATGAAATATGAAGCAAACCTTATG  
GAGTAACCCAGAAATGATGTTGCATTTTTTGTCTTTAGCCTGATAATTGAACTTTCAACA

66617 ATGAAGCAAACCTTTAGGAGTAACCCAGAAATGATGTTGCATTTTTTGTCTTTAGCCTGATA  
ATTGAAACCTTTCAACAATCTCTGGAGTGACTTTTTCTCCTCGAATTGAAACAAGTCTATG  
GCAAAAGAGCTGCATTTTTTTCACAAAAGGGAAGATGGTAACAATGGTCACCTCAAAC  
TTTGGGCTAAATTATATGTACACAGAAATGTTCAAAATCATAGTTTAAATGTGTTTGAA  
AAGGCCACACAATTATACTTTATCTTTCTTAATAATCCTGCAAATCTCTGCCCTGAATC  
[C, T]  
GAAATCTGAAAATGTACTGGCTTGAACAAAATTTGTTTTGTGTGTTAGAGTTATAAATCA  
TTAATCTTTATTTTCGGGTGGTTTACGTTTATGCCAGTTCCTTTATATTTAAATTTCTTGT  
TTTATATATTTTGAATGTCTTTATAGATTTCTTTAAATTCCTTATAGAACCATAATAG  
AAAATCATTACATTTAAATATACCTTACAGCAAAAGCATCCAAATAAGTATAGGGTTTA  
TGTCTTATTTTTCTTTTACGCTGAATACGAATGAGCACAGTGGTGAATTTCTGAAGGGA

66892 ATCCTGCAAATCTCTGCCCTGAATCCGAAATCTGAAAATGTACTGGCTTGAACAAAATTT  
GTTTTGTGTGTTAGAGTTATAAATCATTAACTCTTTATTTTCGGGTGGTTTACGTTTATGCC  
AGTTCCTTTATATTTAAATTTCTGTTTTATATATTTTGAATGTCTTTATAGATTTCTTT  
AAATTTCTTTATAGAACCATTAAATAGAAAATCATTACATTTAAATATACCTTACAGCAA  
AAGCATCCAAATAAGTATAGGGTTTATGTCTTATTTTTCTTTTACGCTGAATACGAATGA  
[G, A]  
CACAGTGGTGGAAATTTCTGAAGGGAAGTGATGAAATTATATTTATTTTACGTGGGCACTTT  
TCCATTTTACCACGTACCATTATTTGGTTCCTGGAGTTATACACTAATTTTACGTATAT  
TACTGTTAAATTACCAACACAAGGCAATTTATTTGAAAGATTCCGTTTATCCTGCCATTG  
CTTTGAAAAGCAGCAGGAAACGAAATCCTTTTACTGTTATGATCAGCTTCTGCAGAGCATCTT  
TGTTTTCTTTTGTCTTTGTTTCTACCTTTTGAATCAGATTCCGTTTATGTCAGGAAGA

67263 CACTGTACCATTATTTGGTTCTGGAGTTATACACTAATTTTACGTATATTACTGTTAAA  
TTACCAACACAAGGCAATTTATTTGAAAGATTCCGTTTATCCTGCCATTGCTTTGAAAAG  
CAGCAGGAAACGAAATCCTTTGACTTGTATCAGCTTCTGCAGAGCATCTTTGTTTTCTTT  
TGTCTTTTGTTCCTACCTTTTGAATCAGATTCCGTTTATGTCAGGAAGACTTCTTGGGA  
CCATTCTTAGTAACCTGAAATTTCTTTTTTAATTGCATGAAGTGGATTGATCATGAGCAA  
[G, A]  
TGATGTGCTTATTTCTCCCTCACTGTTGAATATCTTTGAACTTGCTGTTTTCAATATGGG  
CAGCACAAAGGTGAGAGATACATATTAATAGTAGTATGTATTACTCTTATACATTAGATA  
CCTATATTTAAATGAAAGGCCCAATTTGTAAACATATACATTATATTTCTCTCTTGCCCC  
AAGTTTTAGGAACATGTTAGGATATAGGAGACTTAATTTATAATAATGAGAGCATTTTTT  
TATTTTACTAAAGCCATTTTTATAGTCAACTATCTTTTCTATTTTGTGTATTAGAACTT

67651 ATAGTAGTATGTATTACTCTTATACATTAGATACCTATATTTAAATGAAAGGCCCAATTT  
GTAAACATATACATTATCTCTCTTGTGCCCCAAGTTTATAGGAACATGTTAGGATATAG  
GAGACTTAATTTATAAATGAGAGCATTTTTTATTTTACTAAAGCCATTTTTATAGTC  
AACTATCTTTCTTATTTGTGTGATTAGAACCTAGAAAAATATTTACTAGTTGAAGTTAT  
TATCAGTTTTTAATTTAGTTCTTAAACTCATTTCACTTCTAATAATTTCTGTTATAAATT  
[G, T]  
CCAGCATTTTAAATGAAAATCTAATGATGTAATAGGCATTTTCTTTATTTGAACCTACCTC  
TTTTATTTTCTGAACCAAGAGAAAGATGGACTGGTGTGTTGTTGAAACATTTTTAAAAATG  
TAGTTTTCATTTATATTAGTTATGTTTGATAAATGTCTCAGTATTTTATAATATGATAAG  
CCTGGGATTTCTACTTTTAGGGTTATTTGTACTTTTGAATATATATAAAGTGACAATATT  
AAGGTACATGATCAGCTCTTTCTATTTTTACTCGTAAAAATATGGAATGAATAATTTT

67935 ATTTCTGTTATAAATTGCCAGCATTTTAAATGAAAATCTAATGATGTAATAGGCATTTTCT  
TTATTTGAACCTACCTCTTTTATTTTCTGAACCAAGAGAAAGATGGACTGGTGTGTTG  
AAACATTTTTAAAAATGAGTTTTCATTTATATTAGTTATGTTTGATAAATGTCTCAGTAT  
TTTTATAATATGATAAGCCTGGGATTTCTACTTTTAGGGTTATTTGTACTTTTGAATATA  
TATAAAGTGACAATATTAAGGTACATGATCAGCTCTTTCTATTTTTACTCGTAAAAATTA  
[C, T]

FIGURE 3, page 41 of 42

GGAAATGAATAATTTTGCTAACAACCTTTGAAATTTCAAACCTCTGGAAAATATGAAAATA  
TTCATTGTTTATTATGAATTTAAATTGTAAGGTATGAATGTGATTTGTCTGTACATCTTG  
TATCTTTTCCAAAAAATGATTCTGTATCTTTTGAAAAAAGCCGAGAGTTGAAGATAGTA  
TATTTCTGGTAGTACTGAATATTTTACTTACAGTTTCTATCAAAAAATATATATTTGTTTCT  
AAAATTACTTGTTTTCCAGTTTTTATTTTTTTTAGAGAAAATTCTTAAGTCTCAGTTTCC

69000 TTCAGAAATAACTTATCAGTTATTTCTGTAAGCTTCTTGCTTACCTGGATACCTGACAGG  
TGAGATGGCTGTAGCAGACACTGGCAGTTCCCTGCCCACACACCTGTCCCTGTCCACAGC  
TGCACAAGGCAGCTCTGTGTGCAATTGCCAGCATCTGCTCCTCTGTTCTCAGGGAATCTT  
TGTTAGAAAAATGCTGCCATATTTGTTTCTCACCTATTAGTCTTGTCTCCAGTCAAGAG  
AATAAATTTATGCAAGCAGAGATTGTACTTTACAGTATTTTGTCTTTGAGCTTGGCATT  
[T, G]  
GTTGCATTTGTAAAAATGTGGCATGGCTTCCTCATCCCCAATAGGAACTTTGCCAGCCC  
TTTTGTTCTCATGGAACCTCCTTTTTTGAAAAGAGCACCAAGGAGTAAAAATACTGTGG  
AGGGAGCAACCCTCCTTTGCCATATGCTCTCATTGGGAGACATGTGGAGCAGTCTGAAGT  
CATTTAGGCCACTCTCTGGGAGAGCACATCCTATGATGTTCTCCAGCCTAGCCCCCTTCC  
ACTGTGCTCAAGTCCAAGCTGACCAGCTTTCTGACCACAGTGTAACAAAGATGATTGTC

69134 CTGTGTGCAATTGCCAGCATCTGCTCCTCTGTTCTCAGGGAATCTTTGTTAGAAAAATGC  
TGCCATATTTGTTTCTCACCTATTAGTCTTGTCTCCAGTCAAGAGAATAAATTTATGCA  
AGCAGAGATTGTACTTTACAGTATTTTGTCTTTGAGCTTGGCATTAGGTTGCATTTGTAA  
AAATGTGGCATGGCTTCCTCATCCCCAATAGGAACTTTGCCAGCCCTTTTGTCTCATG  
GAACTTCCTTTTTTGAAAAGAGCACCAAGGAGTAAAAATACTGTGGAGGGAGCAACCCT  
[C, T]  
CTTTGCCATATGCTCTCATTGGGAGACATGTGGAGCAGTCTGAAGTCATTTAGGCCACTG  
TCTGGGAGAGCACATCCTATGATGTTCTCCAGCCTAGCCCCTTCCACTGTGCTCAAGTC  
CAAGCTGACCAGCTTTCTGACCACAGTGTAACAAAGATGATTGTCAAGTGGGCCCCAGAA  
TCCTATACCCAGA